TRB

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2176

Travel Forecasting 2010

VOLUME 2

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

TRANSPORTATION RESEARCH BOARD

OF THE NATIONAL ACADEMIES

TRANSPORTATION RESEARCH RECORDS, which are published throughout the year, consist of collections of papers on specific transportation modes and subject areas. Each Record is classified according to the subscriber category or categories covered by the papers published in that volume. The views expressed in papers published in the Transportation Research Record series are those of the authors and do not necessarily reflect the views of the peer review committee(s), the Transportation Research Board, the National Research Council, or the sponsors of TRB activities. The Transportation Research Board does not endorse products or manufacturers; trade and manufacturers' names may appear in a Record paper only if they are considered essential.

PEER REVIEW OF PAPERS: All papers published in the Transportation Research Record series have been reviewed and accepted for publication through the Transportation Research Board's peer review process established according to procedures approved by the Governing Board of the National Research Council. Papers are refereed by the TRB standing committees identified on page ii of each Record. Reviewers are selected from among committee members and other outside experts. The Transportation Research Board requires a minimum of three reviews; a decision is based on reviewer comments and resultant author revision. For details about the peer review process, see the information on the inside back cover.

THE TRANSPORTATION RESEARCH RECORD PUBLICATION BOARD, comprising a cross section of transportation disciplines and with equal representation from the academic and practitioner communities, assures that the quality of the *Transportation Research Record: Journal of the Transportation Research Board* and the TRB paper peer review process is consistent with that of a peer-reviewed scientific journal. Members from the academic community make decisions on granting tenure; all academic and practitioner members have authored papers published in peer-reviewed journals and all have participated in the TRB peer review process.

TRANSPORTATION RESEARCH BOARD PUBLICATIONS may be ordered directly from the TRB Business Office, through the Internet at www.TRB.org, or by annual subscription through organizational or individual affiliation with TRB. Affiliates and library subscribers are eligible for substantial discounts. For further information, contact the Transportation Research Board Business Office, 500 Fifth Street, NW, Washington, DC 20001 (telephone 202-334-3213; fax 202-334-2519; or e-mail TRBsales@nas.edu).

TRANSPORTATION RESEARCH RECORD PAPERS ONLINE: The TRR Journal Online website provides electronic access to the full text of papers that have been published in the Transportation Research Record series since 1996. The site is updated as new Record papers become available. To search abstracts and to find subscription and pricing information, go to www.TRB.org/TRROnline.

TRANSPORTATION RESEARCH BOARD 2010 EXECUTIVE COMMITTEE*

Chair: Michael R. Morris, Director of Transportation, North Central Texas Council of Governments, Arlington
Vice Chair: Neil J. Pedersen, Administrator, Maryland State Highway Administration, Baltimore
Executive Director: Robert E. Skinner, Jr., Transportation Research Board

J. Barry Barker, Executive Director, Transit Authority of River City, Louisville, Kentucky
Allen D. Biehler, Secretary, Pennsylvania Department of Transportation, Harrisburg
Larry L. Brown, Sr., Executive Director, Mississippi Department of Transportation, Jackson
Deborah H. Butler, Executive Vice President, Planning, and CIO, Norfolk Southern Corporation, Norfolk, Virginia
William A. V. Clark, Professor, Department of Geography, University of California, Los Angeles
Eugene A. Conti, Jr., Secretary of Transportation, North Carolina Department of Transportation, Raleigh
Nicholas J. Garber, Henry L. Kinnier Professor, Department of Civil Engineering, and Director, Center for Transportation Studies, University of Virginia, Charlottesville
Jeffrey W. Hamiel, Executive Director, Metropolitan Airports Commission, Minneapolis, Minnesota
Paula J. Hammond, Secretary, Washington State Department of Transportation, Olympia
Edward A. (Ned) Helme, President, Center for Clean Air Policy, Washington, D.C.
Adib K. Kanafani, Cahill Professor of Civil Engineering, University of California, Berkeley (Past Chair, 2009)
Susan Martinovich, Director, Nevada Department of Transportation, Carson City
Debra L. Miller, Secretary, Kansas Department of Transportation, Topeka (Past Chair, 2008)
Sandra Rosenbloom, Professor of Planning, University of Arizona, Tucson
Tracy L. Rosser, Vice President, Corporate Traffic, Wal-Mart Stores, Inc., Mandeville, Louisiana
Steven T. Scalzo, Chief Operating Officer, Marine Resources Group, Seattle, Washington
Henry G. (Gerry) Schwartz, Jr., Chairman (retired), Jacobs/Sverdrup Civil, Inc., St. Louis, Missouri
Beverly A. Scott, General Manager and Chief Executive Officer, Metropolitan Atlanta Rapid Transit Authority, Atlanta, Georgia
David Seltzer, Principal, Mercator Advisors LLC, Philadelphia, Pennsylvania
Daniel Sperling, Professor of Civil Engineering and Environmental Science and Policy; Director, Institute of Transportation Studies; and Interim Director, Energy Efficiency Center, University of California, Davis
Kirk T. Steudle, Director, Michigan Department of Transportation, Lansing
Douglas W. Stotlar, President and Chief Executive Officer, Con-Way, Inc., Ann Arbor, Michigan
C. Michael Walton, Ernest H. Cockrell Centennial Chair in Engineering, University of Texas, Austin (Past Chair, 1991)

Peter H. Appel, Administrator, Research and Innovative Technology Administration, U.S. Department of Transportation (ex officio)
J. Randolph Babbitt, Administrator, Federal Aviation Administration, U.S. Department of Transportation (ex officio)
Rebecca M. Brewster, President and COO, American Transportation Research Institute, Smyrna, Georgia (ex officio)
George Bugliarello, President Emeritus and University Professor, Polytechnic Institute of New York University, Brooklyn; Foreign Secretary, National Academy of Engineering, Washington, D.C. (ex officio)
Anne S. Ferro, Administrator, Federal Motor Carrier Safety Administration, U.S. Department of Transportation (ex officio)
LeRoy Gishi, Chief, Division of Transportation, Bureau of Indian Affairs, U.S. Department of the Interior, Washington, D.C. (ex officio)
Edward R. Hamberger, President and CEO, Association of American Railroads, Washington, D.C. (ex officio)
John C. Horsley, Executive Director, American Association of State Highway and Transportation Officials, Washington, D.C. (ex officio)
David T. Matsuda, Deputy Administrator, Maritime Administration, U.S. Department of Transportation (ex officio)
Victor M. Mendez, Administrator, Federal Highway Administration, U.S. Department of Transportation (ex officio)
William W. Millar, President, American Public Transportation Association, Washington, D.C. (ex officio) (Past Chair, 1992)
Tara O'Toole, Under Secretary for Science and Technology, U.S. Department of Homeland Security (ex officio)
Robert J. Papp (Adm., U.S. Coast Guard), Commandant, U.S. Coast Guard, U.S. Department of Homeland Security (ex officio)
Cynthia L. Quarterman, Administrator, Pipeline and Hazardous Materials Safety Administration, U.S. Department of Transportation (ex officio)
Peter M. Rogoff, Administrator, Federal Transit Administration, U.S. Department of Transportation (ex officio)
David L. Strickland, Administrator, National Highway Traffic Safety Administration, U.S. Department of Transportation (ex officio)
Joseph C. Szabo, Administrator, Federal Railroad Administration, U.S. Department of Transportation (ex officio)
Polly Trottenberg, Assistant Secretary for Transportation Policy, U.S. Department of Transportation (ex officio)
Robert L. Van Antwerp (Lt. General, U.S. Army), Chief of Engineers and Commanding General, U.S. Army Corps of Engineers, Washington, D.C. (ex officio)

* Membership as of October 2010.

TRB

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2176

Travel Forecasting 2010

VOLUME 2

A Peer-Reviewed Publication

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

Washington, D.C.
2010

www.TRB.org

Transportation Research Record 2176
ISSN 0361-1981
ISBN 978-0-309-16045-2

Subscriber Categories
Highways; public transportation; pedestrians and bicyclists; planning and forecasting; passenger transportation; economics; environment; data and information technology

Printed in the United States of America

TRANSPORTATION RESEARCH RECORD PUBLICATION BOARD

C. Michael Walton, PhD, PE, Ernest H. Cockrell Centennial Chair in Engineering and Professor of Civil Engineering, University of Texas, Austin (Cochair)
Mary Lynn Tischer, PhD, Director, Office of Transportation Policy Studies, Federal Highway Administration, Washington, D.C. (Cochair)
Daniel Brand, SB, SM, PE, Consultant, Lyme, New Hampshire
Mary R. Brooks, BOT, MBA, PhD, William A. Black Chair of Commerce, Dalhousie University, Halifax, Nova Scotia, Canada
Charles E. Howard, Jr., MCRP, Director, Transportation Planning, Puget Sound Regional Council, Seattle, Washington
Thomas J. Kazmierowski, PEng, Manager, Materials Engineering and Research Office, Ontario Ministry of Transportation, Toronto, Canada
Michael D. Meyer, PhD, PE, Frederick R. Dickerson Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta
Sandra Rosenbloom, MA, PhD, Professor of Planning, University of Arizona, Tucson
Kumares C. Sinha, PhD, PE, Olson Distinguished Professor, School of Civil Engineering, Purdue University, West Lafayette, Indiana
L. David Suits, MSCE, Executive Director, North American Geosynthetics Society, Albany, New York

Peer Review of Transportation Research Record 2176

PLANNING AND ENVIRONMENT GROUP
Katherine F. Turnbull, Texas Transportation Institute (Chair)

Transportation System Policy, Planning, and Process Section
Charles E. Howard, Jr., Puget Sound Regional Council (Chair)

Statewide Multimodal Transportation Planning Committee
Patricia G. Hendren, Washington Metropolitan Area Transit Authority (Chair), Adjo A. Amekudzi, Kathleen S. Ames, Jerri L. Bohard, Sonna Lynn Fernandez, Allan Fye, Gregory T. Giaimo, Timothy Harbot, Ashby Johnson, Sandi C. Kohrs, Ronald D. McCready, Michael D. Meyer, Harlan W. Miller, Neil J. Pedersen, Steven M. Pickrell, Camelia Ravanbakht, Peggy A. Reichert, Suzann S. Rhodes, Kyle E. Schneweis, Brian J. Smith, Joan C. Sollenberger, Jack R. Stickel, Montie G. Wade, Rhonda K. Young, Lynn Zanto

Travel Analysis Methods Section
Ram M. Pendyala, Arizona State University (Chair), Barbara J. Arens, Kevin J. Krizek, Eric J. Miller, Srinivas Peeta, Thomas F. Rossi

Transportation Demand Forecasting Committee
Thomas F. Rossi, Cambridge Systematics, Inc. (Chair), Thomas J. Adler, Shlomo Bekhor, Chandra R. Bhat, John L. Bowman, Mark A. Bradley, Stacey G. Bricka, Joseph Castiglione, William Charlton, Elisabetta Cherchi, Mark W. Dunzo, Gregory D. Erhardt, Brian Gardner, Laurie A. Garrow, Gregory T. Giaimo, Konstadinos G. Goulias, Bruce A. Griesenbeck, Jessica Y. Guo, Frank S. Koppelman, T. Keith Lawton, David M. Levinson, Eric J. Miller, Abolfazl Mohammadian, Yasasvi Divakar Popuri, Jeremy Raw, Erik E. Sabina, Venkataraman N. Shankar, Bruce D. Spear, Peter R. Stopher, Harry J. P. Timmermans, Peter S. Vovsha, Joan L. Walker, S. Travis Waller, Edward Weiner

Transportation Planning Applications Committee
Barbara J. Arens, Parsons Brinckerhoff, Inc. (Chair), Elizabeth M. Amar, Eugene L. Bandy, Vincent L. Bernardin, Jr., Mark S. Byram, Elisabetta Cherchi, Michael Clarke, Patrick J. Coleman, Huey P. Dugas, Julie K. P. Dunbar, Karen M. Faussett, Jon D. Fricker, Albert C. Gan, Charles Gorugantula, Kathie M. Haire, Christopher Thomas Hiebert, Leta F. Huntsinger, Nazrul Islam, Dane R. Ismart, Keith L. Killough, David L. Kurth, Richard S. Marshment, Zachary Patterson, Guy Rousseau, Robert G. Schiffer, Peter R. Stopher, Sarah Sun, Dimitrios A. Tsamboulas, Chris Dale Van Slyke, Montie G. Wade, Joan L. Walker, Richard E. Walker, W. Thomas Walker

Peer review is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 2009.

Transportation Research Board Staff
Kimberly M. Fisher, Associate Division Director, Transportation Planning
Freda R. Morgan, Senior Program Associate
Mary Kissi, Senior Program Associate

Publications Office
Cay Butler, Editor; Jennifer Chinworth, Production Editor; Jennifer Correro, Proofreader; Gloria Wright, Manuscript Preparer

Ann E. Petty, Managing Editor; Juanita Green, Production Manager; Phyllis Barber, Publishing Administrator; Jennifer J. Weeks, Manuscript Preparation Manager

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both the Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

TRB SPONSORS*

Transportation Departments of the 50 States and the District of Columbia

Federal Government
U.S. Department of Transportation
Federal Aviation Administration
Federal Highway Administration
Federal Motor Carrier Safety Administration
Federal Railroad Administration
Federal Transit Administration
National Highway Traffic Safety Administration
Research and Innovative Technology Administration

Bureau of Indian Affairs
Science and Technology Directorate,
U.S. Department of Homeland Security
U.S. Army Corps of Engineers
U.S. Coast Guard

Nongovernmental Organizations
American Association of State Highway
and Transportation Officials
American Public Transportation Association
American Transportation Research Institute
Association of American Railroads

---

*As of October 2010.

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2176

Contents

Foreword

The 2010 series of the *Transportation Research Record: Journal of the Transportation Research Board* consists of approximately 900 papers selected from 3,700 submissions after rigorous peer review. The peer review for each paper published in this volume was coordinated by the committee acknowledged at the end of the text; members of the reviewing committees for the papers in this volume are listed on page ii.

Additional information about the *Transportation Research Record: Journal of the Transportation Research Board* series and the peer review process appears on the inside back cover. TRB appreciates the interest shown by authors in offering their papers, and the Board looks forward to future submissions.

Note: Many of the photographs, figures, and tables in this volume have been converted from color to grayscale for printing. The electronic files of the papers, posted on the web at **www.TRB.org/TRROnline**, retain the color versions of photographs, figures, and tables as originally submitted for publication.

Measurement Conversion Factors

To convert from the unit in the first column to the unit in the second column, multiply by the factor in the third column.

Customary Unit	*SI Unit*	*Factor*
Length		
inches	millimeters	25.4
inches	centimeters	2.54
feet	meters	0.305
yards	meters	0.914
miles	kilometers	1.61
Area		
square inches	square millimeters	645.1
square feet	square meters	0.093
square yards	square meters	0.836
acres	hectares	0.405
square miles	square kilometers	2.59
Volume		
gallons	liters	3.785
cubic feet	cubic meters	0.028
cubic yards	cubic meters	0.765
Mass		
ounces	grams	28.35
pounds	kilograms	0.454
short tons	megagrams	0.907
Illumination		
footcandles	lux	10.76
footlamberts	candelas per square meter	3.426
Force and Pressure or Stress		
poundforce	newtons	4.45
poundforce per square inch	kilopascals	6.89

Temperature

To convert Fahrenheit temperature (°F) to Celsius temperature (°C), use the following formula:
°C = (°F – 32)/1.8

SI Unit	*Customary Unit*	*Factor*
Length		
millimeters	inches	0.039
centimeters	inches	0.394
meters	feet	3.281
meters	yards	1.094
kilometers	miles	0.621
Area		
square millimeters	square inches	0.00155
square meters	square feet	10.764
square meters	square yards	1.196
hectares	acres	2.471
square kilometers	square miles	0.386
Volume		
liters	gallons	0.264
cubic meters	cubic feet	35.314
cubic meters	cubic yards	1.308
Mass		
grams	ounces	0.035
kilograms	pounds	2.205
megagrams	short tons	1.102
Illumination		
lux	footcandles	0.093
candelas per square meter	footlamberts	0.292
Force and Pressure or Stress		
newtons	poundforce	0.225
kilopascals	poundforce per square inch	0.145

Temperature

To convert Celsius temperature (°C) to Fahrenheit temperature (°F), use the following formula:
°F = (°C × 1.8) + 32

Abbreviations Used Without Definitions

AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACRP	Airport Cooperative Research Program
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASTM	American Society for Testing and Materials (known by abbreviation only)
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
ITE	Institute of Transportation Engineers
NASA	National Aeronautics and Space Administration
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SHRP	Strategic Highway Research Program
TCRP	Transit Cooperative Research Program
TRB	Transportation Research Board

Integrating an Activity-Based Travel Demand Model with Dynamic Traffic Assignment and Emission Models

Implementation in the Greater Toronto, Canada, Area

Jiang Yang Hao, Marianne Hatzopoulou, and Eric J. Miller

Microsimulation is becoming more popular in transportation research. This research explores the potential of microsimulation by integrating an existing activity-based travel demand model, TASHA, with a dynamic agent-based traffic simulation model, MATSim. Differences in model precisions from the two models are resolved through a series of data conversions, and the models are able to form an iterative process similar to previous modeling frameworks using TASHA and static assignment using Emme/2. The resulting model is then used for light-duty vehicle emission modeling where the traditional average-speed modeling approach is improved by exploiting agent-based traffic simulation results. This improved method of emission modeling is more sensitive to the effect of congestion, and the linkage between individual vehicles and link emissions is preserved. The results have demonstrated the advantages of the microsimulation approach over conventional methodologies that rely heavily on temporal or spatial aggregation. The framework can be improved by further enhancing the sensitivity of TASHA to travel time.

Population growth and urbanization have brought great changes to cities in the past decade. People have more reasons to travel and more ways to do so. As a result, there is increasing interest and need to survey, understand, and predict how people decide to travel; how traveling affects the transportation network; and how a utilized but often stressed transportation network influences the growth of cities and brings changes to the environment.

Modeling the transportation environment and its impact on society requires a variety of models. There are models that focus on specific aspects such as demographic forecasting, car ownership, and trip mode choice. There are also macroscopic frameworks that link multiple submodels to capture the interactions among subsystems. Despite the great variety in scale and scope, current operational models are mostly static and rely heavily on spatial and temporal aggregations to satisfy data limitations and computational constraints. Interactions between models may be hindered by different aggregation techniques, and resolution of the overall system is often determined by the weakest link. Moreover, although such conventional methods provide acceptable systemwide forecasts, they lack the sophistication required for policy analysis. Forecasting people's sensitivity to new policies requires modeling the psychological and behavioral processes that lead to decisions (*1*).

Responding to the growing demand for more efficient, realistic, and policy-sensitive transportation models, microsimulation that explicitly models the dynamic behavior of each individual agent is becoming a more popular approach to transportation and land use modeling practices (*2–4*). The disaggregate nature of this technique makes it capable of providing a much finer resolution of the system state. A fully integrated microsimulation model for travel demand and traffic assignment is capable of providing detailed agent-based information for the development and evaluation of transport policies.

In terms of policy evaluation, the assessment of environmental performance, in particular, has grown in importance mostly because of increasing concern about environmental preservation, resource consumption, limiting greenhouse gas emissions, and human exposure to air pollution. The latest Environmental Protection Agency report on air-quality trends indicates that highway vehicles alone account for 52% of carbon monoxide (CO) emissions, 33% of nitrogen oxide (NO_x) emissions, and 23% of volatile organic compound emissions (*5*). Increased car use induced by urbanization and population growth has led to more attention being placed on vehicle emissions. More methods for measuring and predicting vehicle emissions have been developed in recent years in response to the increasing number of emission regulations. Estimating emission levels for the current fleet can be achieved through various inspection and surveying techniques, which are almost always more reliable than estimates from emission models (*6*). Nevertheless, emission models are useful, especially for forecasting and comparing emission levels under different policy scenarios and determining whether the proposed project conforms to regulatory requirements.

Considerable research into integrating emission models with travel demand models as well as integrated land use and transport models has been recently conducted and implemented in cities such as Kyoto, Japan; Portland, Oregon; Sacramento, California; Mexico City, Mexico; and Dortmund, Germany (*7–10*). Most efforts focus on light-duty vehicles and are integrated with static emission models, which generate average hourly or daily emissions for the area of interest, aggregated over different vehicle types, ages, and other characteristics. The research reported here has two main objectives. The first objective is to develop an agent-based travel demand and traffic assignment modeling framework by integrating existing

J. Y. Hao and M. Hatzopoulou, Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario M5S 1A4, Canada. E. J. Miller, Department of Civil Engineering and Cities Center, University of Toronto, 455 Spadina Avenue, Suite 400, Toronto, Ontario M5S 2G8, Canada. Corresponding author: E. J. Miller, miller@civ.utoronto.ca.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 1–13.
DOI: 10.3141/2176-01

agent-based software, namely, TASHA and MATSim. The second objective is to use this newly integrated model to obtain a finer resolution for vehicle emission modeling, thus demonstrating the advantages of the microsimulation approach over conventional methodologies that rely heavily on temporal or spatial aggregation. The modeling framework is applied to the greater Toronto area (GTA).

TASHA–MATSim INTEGRATION

This section describes the integration of TASHA and MATSim, beginning with a brief description of the two models, followed by issues encountered during the data conversion process. Finally, the performance of this integrated microsimulation framework is analyzed by comparing it with a previous modeling framework using TASHA and Emme/2.

Description of Models

TASHA

TASHA is a microsimulation model for forecasting travel demand. It is designed to improve on conventional four-stage models. The model operates based on three hypotheses:

> First, a fundamental assumption is that scheduling is an event-driven, sequential process, in which individual episodes are provisionally scheduled as they arise out of personal and household projects.
>
> Given this, it follows that, in general, activity-travel is not an optimizing procedure.
>
> A third fundamental assumption is that travel mode choice (and the associated allocation of household vehicles for individual person travel, as required) is inherent in the activity scheduling process. (*11*, p. 115)

The main components of TASHA are the scheduler and mode choice model. An additional trip processing module, named Merge_tt, links these two components (Figure 1). Its purpose is to build trip chains based on the feasibility of individual trips. The conceptual design of TASHA is shown in Figure 2 and is described in detail by Miller and Roorda (*11*). Although the model is agent based, it still follows a sequential decision-making process similar to that of a conventional four-stage model. The program generates activities for each person based on observed joint probability distributions, predicts activity locations based on a series of entropy models, schedules activities following a set of rules, and assigns mode through a random utility tour-based mode choice model. It has been validated with 2001 trip diary data (*12*). The validation exercise uses TASHA with parameters estimated using 1996 Transportation Tomorrow Survey data to forecast 2001 activities. The results demonstrate that TASHA is capable of accurately predicting distributions for activity frequencies, start times, durations, and trip distances.

MATSim

MATSim (also referred to as MATSim-T) is being developed jointly at Technische Universität Berlin; Eidgenossische Technische Hochschule Zurich, Switzerland; and Centre National de la Recherche Scientifique Lyon, France. It consists of a variety of microsimulation tools for modeling travel demand and traffic flow. MATSim is a fast agent-based simulation designed to handle large networks and millions of agents (*13*). Large-scale scenarios have been tested in Zurich, Berlin, and other cities. To manage the wide range of submodels and algorithms, MATSim adopts a modular approach with a standardized data format, so that new modeling and analysis modules can be integrated into the framework in a plug-and-play style.

This project focuses on the Iterative Demand Optimization Process–Evolutionary Algorithm (MATSim-EA) within MATSim. Starting with an initial travel demand, the simulation is an iterative process between travel activity rescheduling and traffic assignment. Schedules can be modified through three strategies: a dynamic Djikstra router that changes the traveling route; the time allocation mutator, which shifts activity start times by a specified amount; and rescheduling the whole activity with a household scheduler module named planomat. The scheduling and rescheduling processes in MATSim and TASHA differ in the way schedules are evaluated and in how new schedules are generated.

Unlike TASHA, which uses a set of rules to assess feasibility of the schedules, MATSim evaluates each schedule according to a utility-based scoring function. The utility gain for performing an activity is logarithmically related to its duration to reflect the diminishing marginal utility gains as the duration increases. The disutility for travel is linearly related to the overall travel time. Delays and other complications in travel are indirectly represented as opportunity loss for not performing activities. A detailed explanation of this concept and methodology can be found in the original paper by Charypar et al. (*14*). A schedule needs to be modified or replaced when it is infeasible as viewed by TASHA or from the MATSim's perspective has a low performance score. MATSim comes with a rescheduling module called planomat. Compared with the scheduler in TASHA, planomat uses a very different concept for generating new schedules. The TASHA scheduler replaces infeasible activities within a schedule by drawing a new one from activity frequency distributions, whereas planomat uses a genetic algorithm to create a new schedule based on existing ones. After multiple iterations, schedules from TASHA still represent the initial activity frequency distributions. MATSim schedules may drift toward a set of more optimal schedules that "make best use of time" but the definition of optimal depends on the specifications of the scoring functions.

For traffic assignment, MATSim uses stochastic, queue-based agent traffic simulation with a set of very simple rules. Vehicles enter and exit links on a first-in, first-out basis. Vehicles cannot exit unless they have spent enough time on the link to travel its full length. Vehicles cannot enter a link whose storage capacity has been reached and cannot leave a link that reaches flow capacity (*13*). The result is a very fast simulation that produces outputs meaningful and useful for most transport planning purposes. No lane changes, signal schedules, or turn restrictions are modeled. However, it may be possible to model some of these scenarios indirectly by modifying input network or other simulation parameters.

Model Integration

The previous modeling framework had TASHA interacting with Emme/2 to obtain network information through static traffic assignment (*11*). TASHA outputs were converted into origin–destination flows and sent into Emme/2. In return, Emme/2 produced interzonal travel times as input for the TASHA scheduler. For nondrive modes, Emme/2 also updated various levels of service information. For activities that become infeasible because of travel time constraints, the TASHA scheduler would draw a new activity from trip generation distributions. The process is iterated three times.

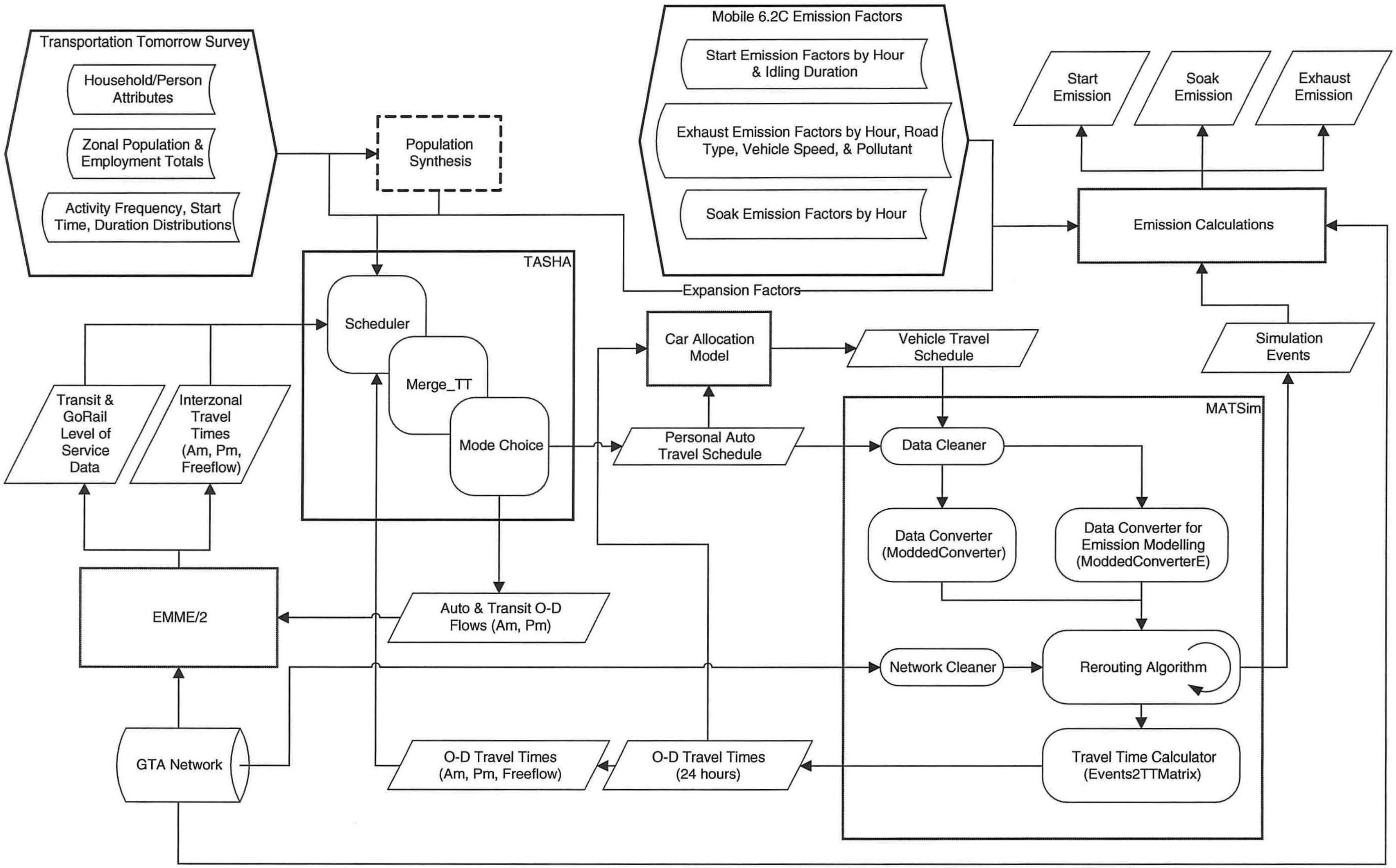


FIGURE 1 Complete system flowchart.

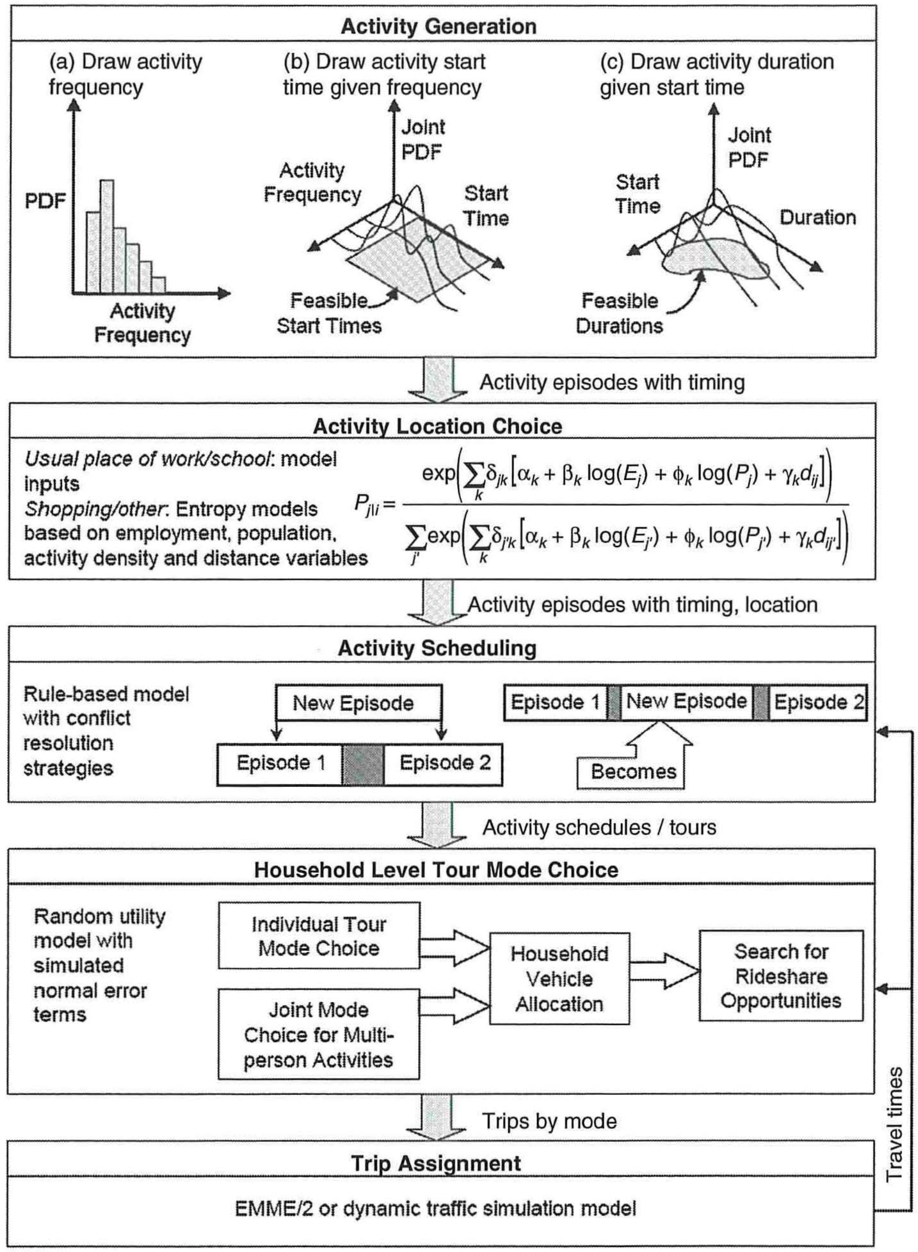


FIGURE 2 TASHA model flowchart (*11*).

Currently, TASHA has been integrated with MATSim, whereby TASHA generates personal tour information and converts it into a format that can be read by MATSim. MATSim then simulates the tours and attempts different routes to minimize and stabilize individual travel times. Once the simulation is complete, interzonal travel times are extracted from the simulation results and fed back into TASHA. As the current version of MATSim can handle only automobile trips, TASHA still has to rely on Emme/2 for non-drive-related data such as transit travel times and level-of-service data (Figure 1).

Network and schedule information needs to be converted into XML file format readable by MATSim. Preparation of the network file is relatively simple as most node and link variables required by MATSim are readily available in Emme/2. For converting the schedules, the main challenge is that information from TASHA is trip based, whereas MATSim requires an event-based schedule. Therefore, each schedule in TASHA needs to be categorized into data associated with the start and end of each activity. In addition, the coordinates of each activity need to be generated based on the coordinates of zone centroids because TASHA locates activities at the centroids of traffic analysis zones. This procedure helps disperse the population so that activities start and end on different links without overloading particular links, while remaining within proximity of the zonal centroid. A built-in algorithm in MATSim was used for this task.

Simulation Configuration

MATSim simulation parameters can be modified within a configuration file. The parameters are organized into categories that manage

different aspects of the model, such as input–output paths, number of iterations, network capacity modifiers, and values for scoring function parameters. For this project, MATSim was configured to perform traffic assignment for 50 iterations before computing a travel time matrix to be fed back into TASHA. The probability of a person attempting a new route was set to be 20%. For the remaining 80% of the time, the person would choose a previously attempted route based on its score performance. With this configuration, MATSim iteration took about 8 min to complete, leading to a total time of about 6 h for a MATSim run with 50 iterations.

TASHA–MATSim Modeling Results

The primary output of MATSim is a list of events. An event is defined as a change in the agent's status, such as starting or ending an activity or entering or exiting a link. MATSim includes various built-in modules for analyzing this event file. For example, Figure 3 shows the score statistics obtained during the 50 MATSim iterations as well as a visualization of the network plotted with traffic information extracted from the event file.

Each person can remember up to five schedules, and, for each new schedule evaluated, the one that has the lowest score will be removed from the choice set. The score plot shows for each iteration the average of people's best, worst, and average scores as well as the one chosen for that iteration. Even after 50 iterations, the average of people's best scores has not completely converged. Complete convergence could be achieved after 300 iterations.

The score first improved quickly, but then the increase after 40 iterations became very small. Because only the rerouting algorithm was used, there were very limited possibilities for further improvement after the first few trials. Therefore, it is sufficient to use 50 iterations for this research.

The worst score decreased during earlier iterations as people added more schedules to their choice set. Once most people had attempted five different schedules, the average of their worst scores started to improve as they started to remove the schedule with the lowest score. However, again because of the limited ways to improve their schedules, the average of people's lowest scores did not improve much before converging.

The road network snapshot of the city of Toronto, Canada, in Figure 3 was obtained with the built-in visualizer OTFVis, which uses the event file to produce snapshots of the system at specified intervals. Agents are represented by car icons on the links. The color of the icons ranges from green to yellow to red to indicate free-flow traffic, mild congestion, and saturated links, respectively. The spacing between icons on the links indicates travel speed. The snapshot shows the network at 8:35 a.m. A large number of links are congested, including many arterial roads and large segments of the freeway leading into downtown Toronto.

Additional statistics on trip length, duration, and interzonal travel time were also extracted (Figure 4). Average trip length starts at a peak of 25 km at 4 a.m. and gradually decreases throughout the day until reaching the afternoon peak, where another high of 17 km can be observed. This process suggests that the few people who travel very early in the morning do so because their destinations are far away. The afternoon peak also reflects long return home trips from downtown Toronto to suburban areas. However, average trip length in the afternoon is reduced by the large number of short shopping and leisure trips. A similar trend can be observed for the average trip duration plot, with the exception of long trips in early morning having much shorter duration due to less congestion. Average zone-to-zone travel time peaks at 53 min in the morning and 50 min in the afternoon. The averages for the peak periods are even smaller because of the sharp decrease in travel time in adjacent hours. These average travel times are shorter than those previously estimated by Emme/2, which predicts 55 and 53 min for the morning and afternoon peak periods, respectively. The off-peak travel time, however, is higher, averaging around 41 min compared with the Emme/2 average of 37 min. This difference is expected because the free-flow travel time in Emme/2 is computed by using the actual free-flow speed, which should be the minimum possible travel time between zones. More details on the performance comparison between Emme/2 and MATSim is described by Gao (*15*). Overall, MATSim and Emme/2 produce similar results in terms of trip distance and link volumes, but MATSim performs better with respect to travel time and link speed because of its greater sensitivity to congestion.

EMISSION MODELING USING TASHA–MATSim

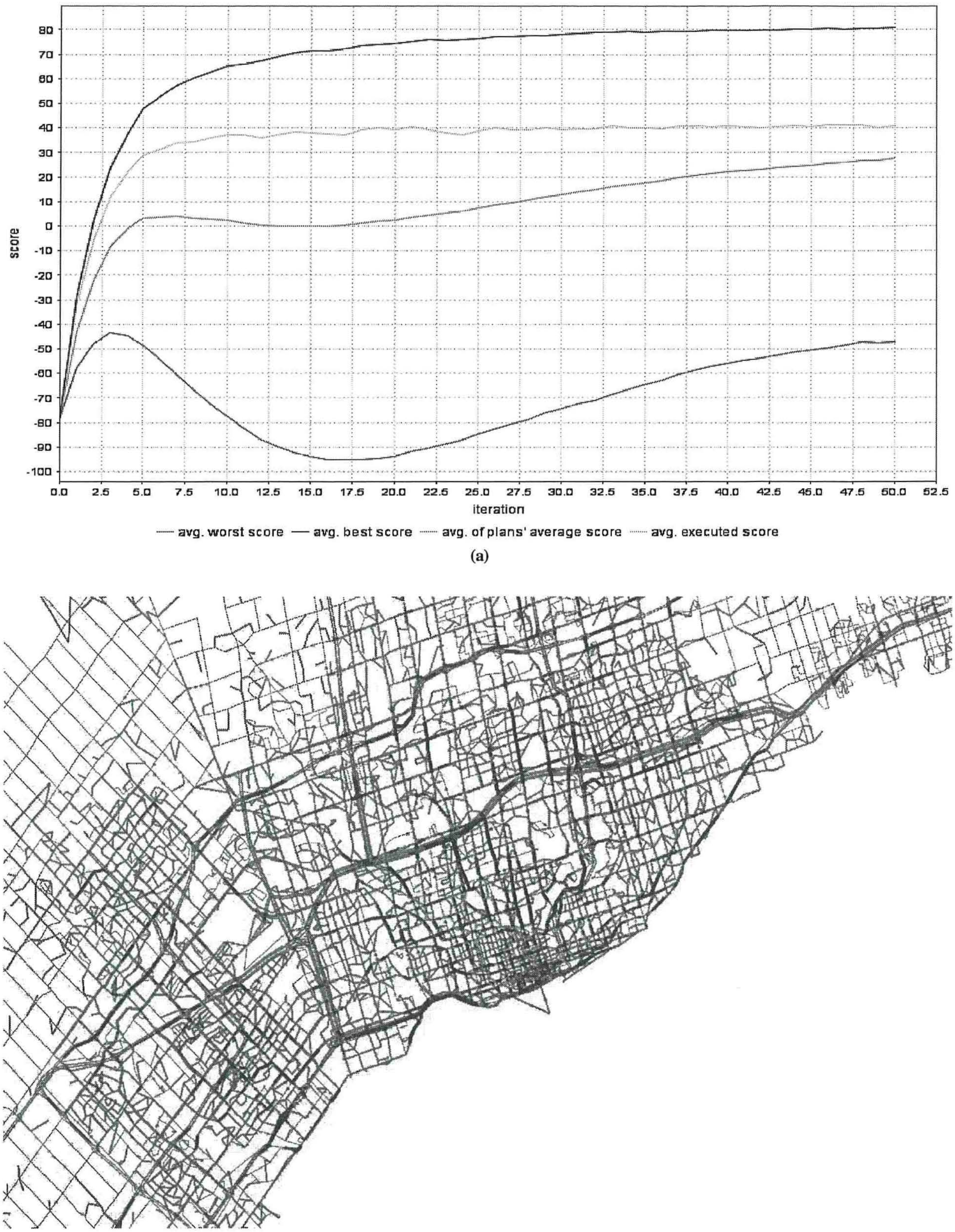
By exploiting the agent-based output generated by the TASHA–MATSim model, emissions on the network are calculated without losing linkage to each household agent. Thus, emissions on the network can be tracked back to those who are producing them, allowing for analysis by household location or various personal attributes. Three major steps are involved: converting travel schedules into car usage schedules by using a car allocation model, constructing look-up tables for different types of emission factors, and using a calculation program to estimate automobile emissions based on vehicle usage information and emission factors.

Car Allocation Model

The purpose of this model is to add the missing link between activity schedules and vehicle emission estimations: vehicle usage information. Inputs for the model are derived from personal trip-chain information predicted by the mode choice submodel within TASHA. Because emissions are generated whenever a vehicle is being used, the inputs include all automobile components of the trip chains. The model assigns a vehicle to each automobile trip within each person's activity chain, taking into account the availability of vehicles as well as scheduling conflicts between different people in the same household. Because of lack of data on vehicle types and people's preferences when choosing a vehicle, the model assumes that all vehicles are identical and assigns them on a first-come, first-serve basis. People's preferences have been simplified to a basic rule that makes people prefer the first car assigned to them. A detailed description of the car allocation model is presented elsewhere (*6*).

MATSim Simulation

Outputs from the car allocation model were first integrated into TASHA schedules by linking the household, person, and trip identification (ID). The schedule converter in MATSim was modified to use household-person-car as the agent ID instead of household-person. As a result, the MATSim simulation now treats each person-car combination as an agent. With this special ID, emission calculation results can be accumulated by person or by car, which is useful when emission factors for different vehicle types are available. Other



(a)

(b)

FIGURE 3 MATSim simulation results: (*a*) score statistics and (*b*) network visualization.

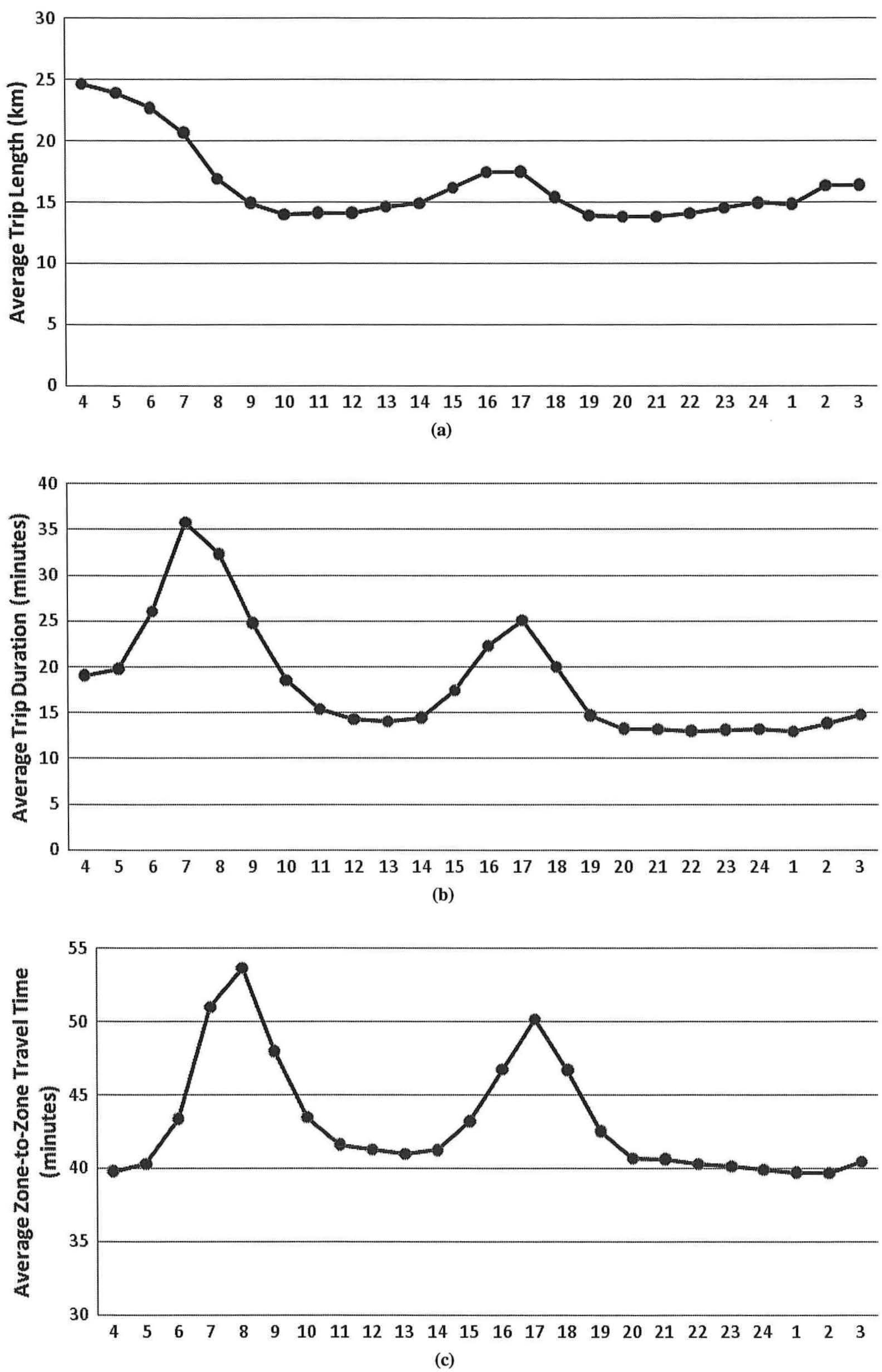


FIGURE 4 Hourly statistics based on MATSim outputs: (*a*) trip length, (*b*) trip duration, and (*c*) zone-to-zone travel time.

than the modification made to the agent ID, the simulation is identical to the previous TASHA–MATSim runs.

Development of Emission Factors

Four types of emissions were estimated: exhaust, idling, start, and hot soak. Exhaust and idling emissions account for vehicle emissions on the roads, whereas start and hot-soak emissions take place at activity locations. These locations occur at particular points on the road network. To estimate emission factors (EFs) for the four types of emissions, three look-up tables for EFs were generated by using Mobile6.2C: exhaust, start, and hot-soak EFs.

Input data for Mobile6.2C is divided into six categories: external conditions, environmental effects on air conditioning, vehicle fleet characteristics, fuel characteristics, state programs, and vehicle activity (*16*). Besides vehicle activity, input data reflecting the other five main categories are derived from various Canadian sources, including Environment Canada (*17, 18*) and the Ontario Drive Clean Program (*19, 20*). The benefit of using GTA-specific distributions was demonstrated in an earlier study where EFs derived from Mobile6.2 default distributions and GTA-specific distributions were compared (*6*). The default Mobile6.2 scenario underestimated GTA emissions. Vehicle activity inputs for Mobile6.2 include vehicle miles traveled (VMT) by vehicle class, VMT by facility, VMT by hour, VMT by speed distribution, starts per day, distribution of vehicle starts during the day, weekday trip length distribution, soak distribution, hot-soak activity, and diurnal soak activity. Vehicle activity distributions were not directly input into Mobile6.2C as the model was used to generate look-up tables rather than average EFs representing those distributions.

For the purpose of developing speed–roadway-type EF look-up tables for exhaust emissions (in g/VMT), the average speed command was used whereby one specific speed attributed to a specific roadway type was input as a scenario. A total of 60 speed–roadway-type scenarios were modeled for each hour of the day, emission type, and pollutant. Speed categories amount to 15 and include 2.5, 5, 7.5, and 10 to ≥65 mph in 5-mph increments. There are four roadway-type categories: freeway, arterial, local, and ramp. Pollutants include hydrocarbons (HC), CO, NO_x, and carbon dioxide (CO_2). Idling emissions are considered to be the same as those occurring at 2.5 mph for the idling duration.

A look-up table for start EFs (in g/start) was also developed. Start emissions depend on the amount of time an engine was turned off before it is restarted (i.e., the soak duration). The look-up table for start EFs was developed for 70 soak duration categories (<1 min, 1 to 30 min in 1-min increments, 30 to 60 min in 2-min increments, and 60 to 720 min in 30-min increments). Start EFs were also developed for the 24 h of the day and for three pollutants: HC, CO, and NO_x.

Finally, a look-up table for HC hot-soak EFs (in g/min) was generated. Hot-soak emissions depend on the length of the soak after an engine shut-down and they typically become negligible within an hour. Thus, the look-up table for hot-soak EFs was developed for 60 1-min soak duration categories.

Emission Calculations

Figure 5 presents the program flowchart and detailed calculation steps performed by the custom-made emissions calculator. The program was developed in C#. In the flowchart, processes above the dashed line are called sequentially by the program class. To reduce the amount of postprocessing, the program accumulates emissions for the network, links and roads on the network, and households as soon as emissions are calculated. The program is also flexible enough that it is capable of outputting each calculation result with associated link or person attributes.

Results and Analysis

Emissions were computed and analyzed through various dimensions, including total emissions in the GTA aggregated by type of emission and pollutant, emissions per trip (time, activity type, trip purpose), emissions per person (age, occupation, gender) and household (household location, dwelling type, number of cars), and emissions by roadway link.

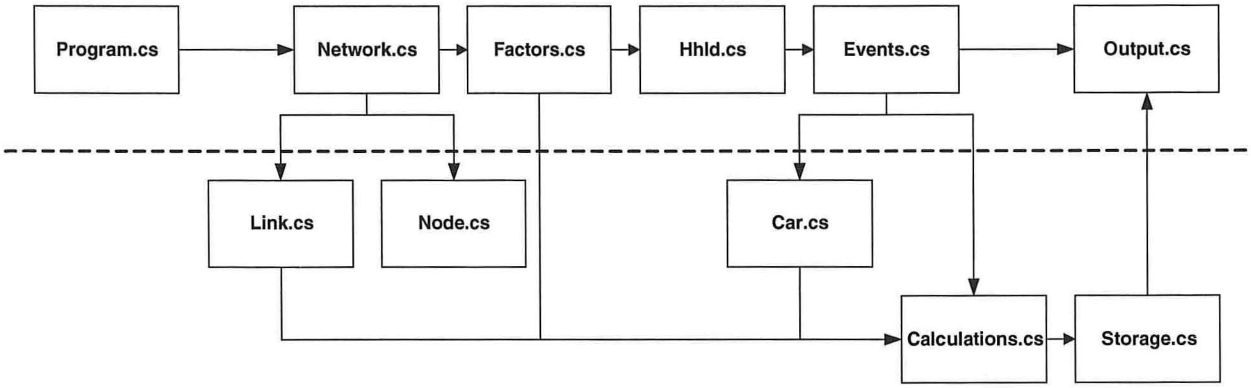
Total Emissions

Total daily emissions for the GTA amounted to 63.95 tons of HC (as volatile organic compounds), 1,362 tons of CO, 75.62 tons of NO_x, and 25.81 kilotons of CO_2. Exhaust emissions have the largest percentage share, followed by start, hot-soak, and idling emissions. As shown in Figure 6, estimates of the emissions on the road (exhaust and idling) are significantly higher than the results obtained using average link speed emissions from Emme/2 (*6*). The difference—about 19% for HC, 13% for CO, 9% for NO_x, and 7% for CO_2—is mainly due to the increased sensitivity to congestion obtained by separately calculating exhaust and idling emissions using MATSim events. All parts of the trips on the network are accounted for including the intrazonal portion not captured by Emme/2. The large difference in HC emissions is also due to the relatively large exhaust EF at low speed, making the effect of idling more apparent.

Hourly Emissions on the Network

Results from the emission calculation program were aggregated by hour. Start emissions have a very short duration so the program assumes they occur at the time of departure. Soak begins at the time of arrival and the program is able to handle properly the case in which the soak period extends to the next hour. For exhaust and idling emissions, the hour is determined by the time a vehicle enters a link. This estimation is considered to be sufficient, as the time spent on most links is much shorter than 1 h, and the difference between EFs from adjacent hours is minimal.

Figure 6 shows the hourly HC emissions in two graphs: the first one shows the total emissions by hour and the second one shows the percentage share of different emission types. Total emissions peak at 8 a.m. and 5 p.m., producing 6.29 and 6.57 tons of HC, respectively. There is a drastic decrease in total emissions in the middle of the day between the two peak periods. This decrease is mostly due to the decrease in exhaust and idling emissions. There are very small idling emissions between the peak periods when there is less congestion. Although start and soak emissions also peak during the peak hours, they remain significant throughout the day. Start emissions are slightly higher in the morning than in the afternoon, because the first start of the day produces more emissions than later starts. Figure 6 shows that, although most emissions are caused by vehicles cruising on the road, idling emissions are very significant at peak hours



Event Type	Variables Updated / Recorded	Values Calculated*
Departure	If this is the first departure**: - record all event attributes***	If previous event is **Arrival**: - soak duration = **Departure** time – **Arrival** time - calculate soak emission at time of **Arrival**: • duration is capped at 1 hour • if soak finished in the next hour, separate soak duration into two components • look up soak emission factors (grams/hour) for each component of the duration ➢ Soak Emission = ∑ emission factor x duration - Calculate start emission at time of **Departure**: • check duration category • obtain hour of day using **Departure** time • look up start emission factor in grams ➢ Start Emission = emission factor
Enter Link	- record all event attributes	
Leave Link		If previous event is **Enter Link**: - calculate exhaust emission at time of **Enter Link**: • free-flow time = link length / free-flow speed • actual time = **Leave Link** time - **Enter Link** time • idling time = actual time – free-flow time • check speed category • obtain hour of day using **Enter Link** time • look up link type (freeway, arterial, local, or ramp) • look up exhaust emission factor in grams/mile • idling emission factor = exhaust emission factor at 2.5 mph ➢ Exhaust Emission = emission factor x link length ➢ Idling Emission = emission factor x idling time x 2.5 mph
Arrival	- update all event attributes	

Notes:
* All values calculated are then amplified by the household expansion factor.
** Since start emission calculation requires information on the previous arrival event, each person's first departure is recorded and then "playback" at the end to calculate overnight soak duration.
*** Event attributes include: Time, CarID, LinkID, Event Type

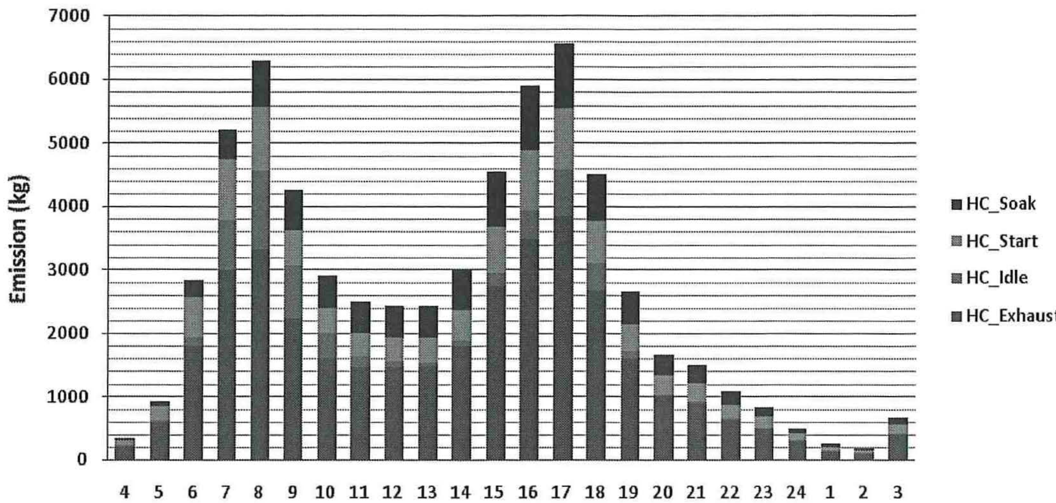
FIGURE 5 Emission calculator program flowchart and calculation steps.

and account for about 20% of morning peak and 12% of afternoon peak emissions.

Household Emissions

The new agent-based output from TASHA–MATSim allows emissions on the network to be calculated without losing linkage to each household agent. As a result, emissions on the network can be tracked back to those producing them, allowing for analysis by household location or various personal attributes. For example, Figure 7 examines the relationship between household emissions and the number of vehicles. When there are five or fewer vehicles in the household, there is a strong correlation. The relationship is not linear and an additional vehicle provides a smaller increase in household emissions. A correlation cannot be observed when the household owns a large

	Emission in tons				% Difference from EMME/2			
	HC	CO	NO_x	CO_2	HC	CO	NO_x	CO_2
EMME2 Avg. Speed	36.33	1,103	63.81	24,120				
MATSim Avg. Speed	42.30	1,180	68.10	25,800	16.43	6.96	6.72	6.97
MATSim Exhaust + Idling	43.16	1,248	69.40	25,810	18.81	13.12	8.76	7.01



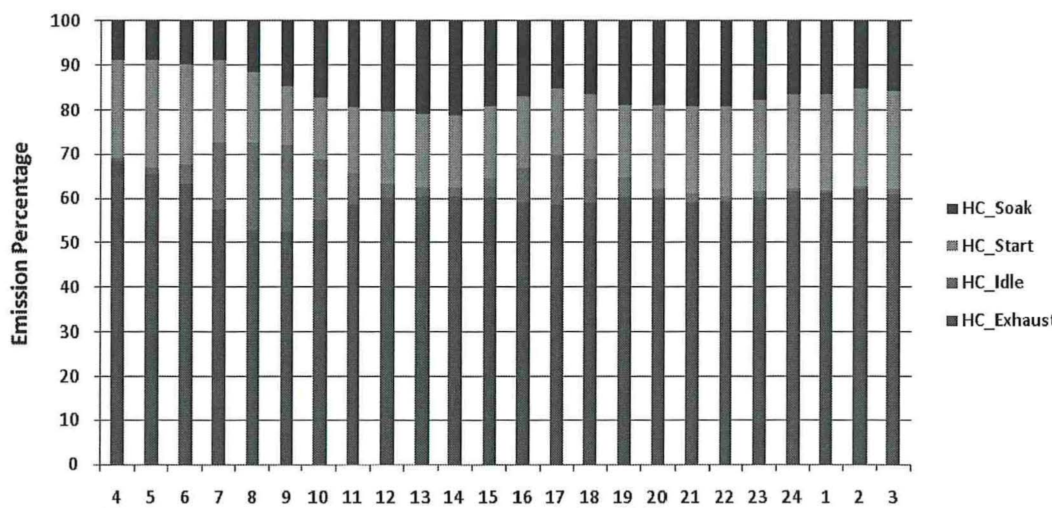


FIGURE 6 Emission estimation results for GTA.

number of vehicles, partially because of lack of data and also probably because of underutilization of the vehicles.

Household location is also a contributing factor in the amount of emissions produced. People living in suburban areas are more likely to drive and to drive further distances. ArcGIS software was used to link the GTA zone boundary file with household emission results. The data used for producing the graph were weighted by household expansion factors, and emissions per household were plotted instead of zonal totals to eliminate bias due to different zone sizes. As shown in Figure 7, households in zones farther from downtown produce more emission in general. Most households located within the highway ring produce fewer than 50 g of HC a day. Newly developed residential areas in the northeast have higher emissions per household.

Link Emissions

Link emissions were also plotted in ArcGIS. The emission amount was divided by link length to remove the bias toward longer links and produce smoother color transition between links. The color scale was generated with data from a particular hour of the day (8 a.m. for exhaust and idling, 5 a.m. for start and soak) and then used for all other hours. The division of ranges was based on Jenk's natural breaks suggested by ArcGIS. This method classifies data by minimizing the sum-of-squared differences between class members and class means. The ranges obtained were further modified through rounding and slightly decreasing the limit of higher ranges to obtain better color distribution.

Figure 8 shows hourly exhaust emission plots for HC. The peak in link exhaust emissions corresponds to the peak in travel demand.

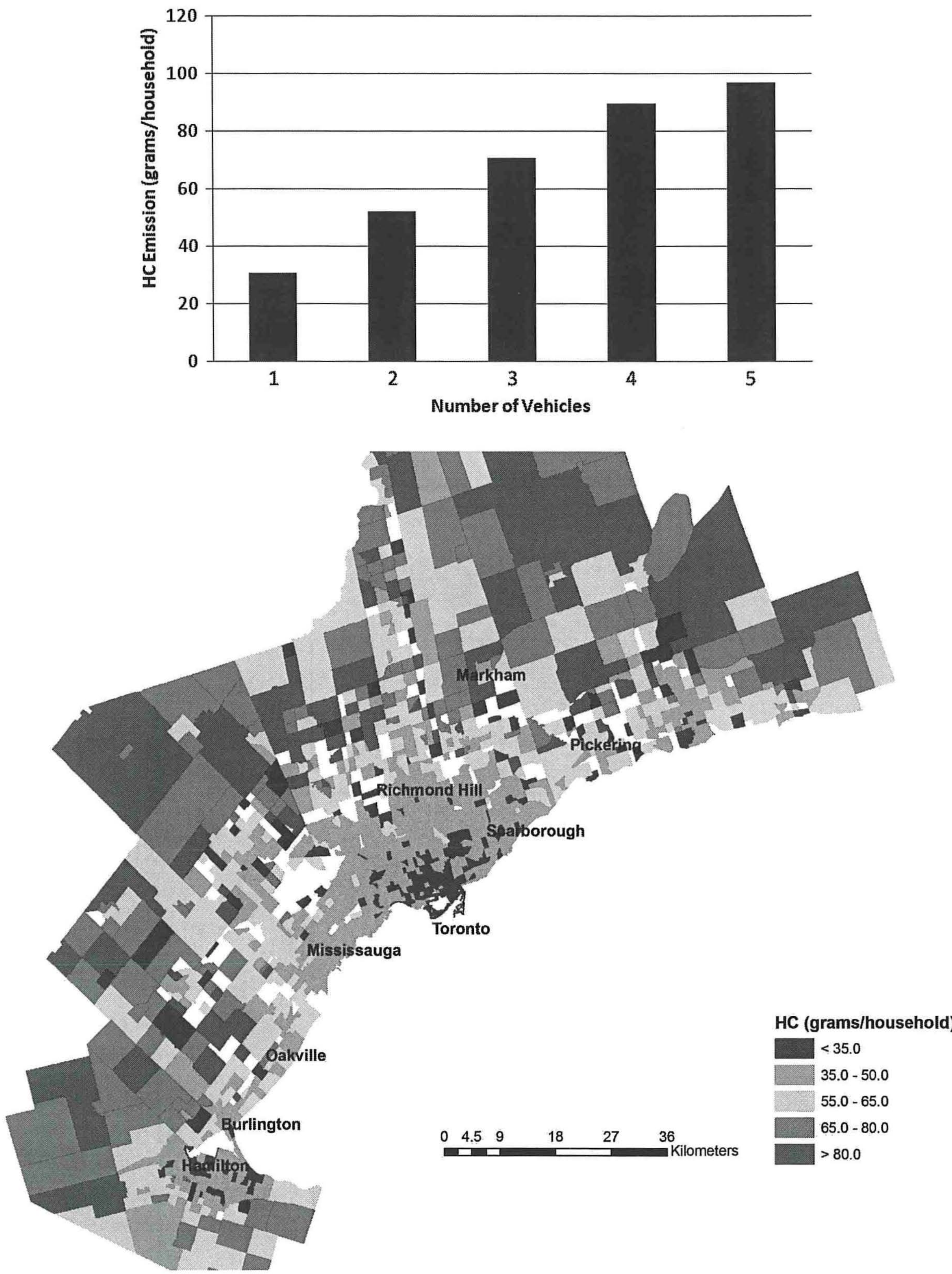


FIGURE 7 Analysis of household emissions.

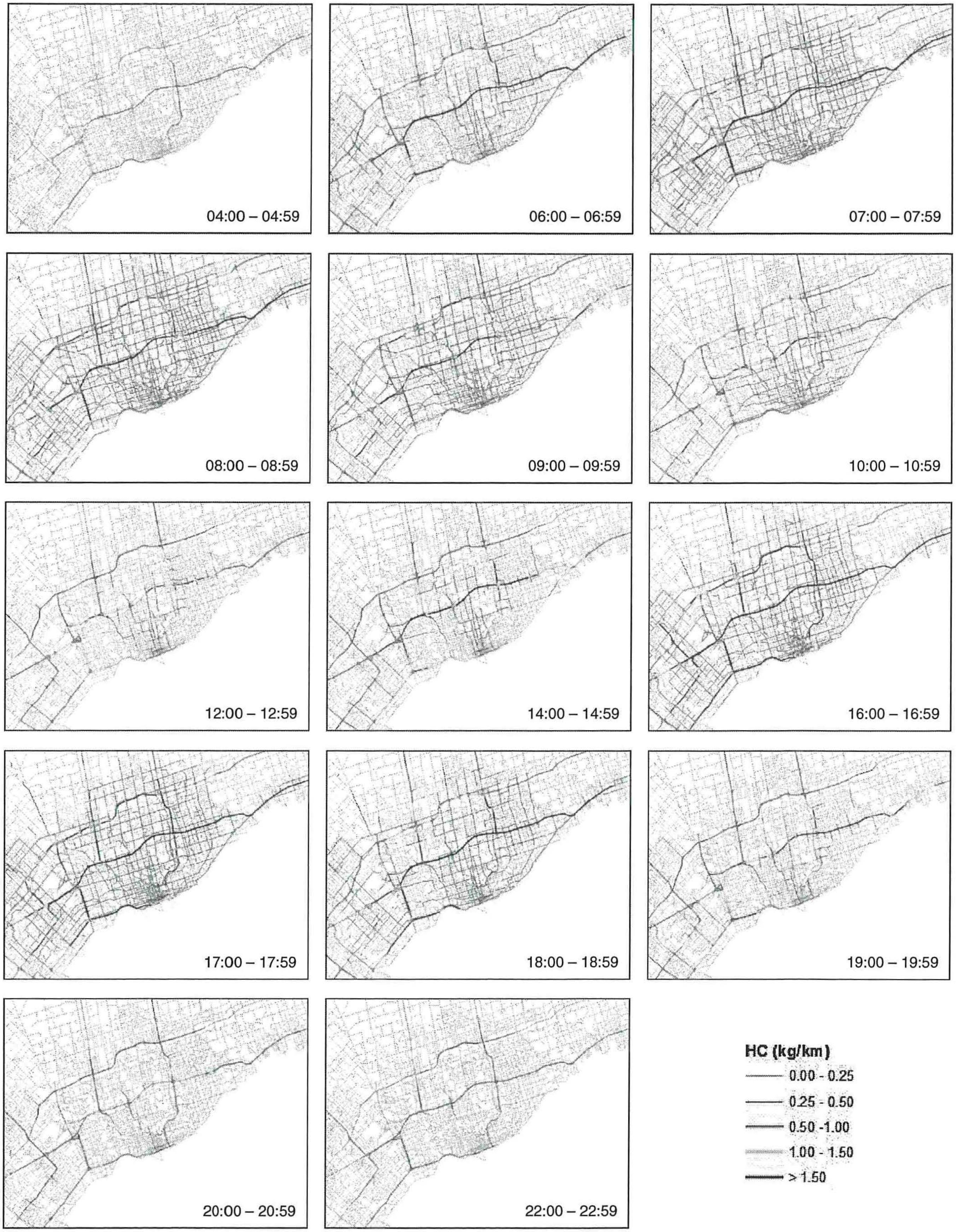


FIGURE 8 Hourly link HC emissions.

In general, highways and arterial roads have much higher exhaust emissions than local streets. The highest emitting arterials are mainly in the city of Toronto and extend north toward York, in addition to Mississauga in the Peel region (west of Toronto) and Hamilton. These areas are major urban centers; yet, emissions from links within the city of Toronto are by far higher because of the large number of vehicles entering the city during the morning peak as well as low traffic speeds on the city's network. Comparing these link emission plots with the household emission plots in Figure 7 shows a large portion of emissions occurring within the city of Toronto, whereas the heavy polluters live in suburban areas.

CONCLUSION

This research explores the potential of microsimulation models by experimenting with the integration of TASHA and MATSim and its application in vehicle emission modeling. The results indicate that the TASHA–MATSim framework is capable of producing meaningful outputs similar to that of TASHA–Emme/2, without losing linkage to the agents. The results, however, may be too similar to TASHA–Emme/2 results, implying that the current version of TASHA is insensitive to zonal travel time variations.

On the emission modeling side, the emission calculation program is an improved average-speed model that adds congestion sensitivity to the calculations. Based on MATSim simulation events, the emission estimates are higher than those obtained through Emme/2. The amount of idling emissions, especially during peak periods, indicates that the model is sensitive to congestion. The agent-based output from MATSim also allows emission analysis to be based on various person and household attributes. The results confirmed that, although most emissions occur within the city of Toronto and other regional centers, heavy polluters live in suburban areas. This finding demonstrates that the new modeling framework is a promising tool that provides a better understanding of how people's behavior affects the system. Future work will implement this framework within a large-scale integrated microsimulation framework, thus fully using the linkage between agents and the system.

REFERENCES

1. Doherty, S. T. Activity Scheduling Process Approach to Understanding Travel Behavior. Presented at 79th Annual Meeting of the Transportation Research Board, Washington, D.C., 2000.
2. Doherty, S. T., and K. W. Axhausen. The Development of a Unified Modeling Framework for the Household Activity-Travel Scheduling Process. Presented at 4th NECTAR Conference, Tel Aviv, Israel, April 1998.
3. Timmermans, H. J. P. Models of Activity Scheduling Behavior. *Proc., 2nd Aachen Colloquium, 'Mobility and City,'* Aachen, Germany, 2001, pp. 33–47.
4. Arentze, T. A., and H. J. P. Timmermans. *Albatross: A Learning Based Transportation Oriented Simulation System.* European Institute of Retailing and Services Studies, Eindhoven, Netherlands, 2000.
5. *National Air Quality—Status and Trends Through 2007.* Office of Air Quality Planning and Standards Research, U.S. Environmental Protection Agency, Washington, D.C., 2008. www.epa.gov/air/airtrends/2008/. Accessed April 2009.
6. Hatzopoulou, M., E. J. Miller, and B. Santos. Integrating Vehicle Emission Modeling with Activity-Based Travel Demand Modeling: Case Study of the Greater Toronto, Canada, Area. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2011,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 29–39.
7. Lautso, K., K. Spiekermann, M. Wegener, I. Sheppard, P. Steadman, A. Martino, R. Domingo, and S. Gayda. *PROPOLIS: Planning and Research of Policies for Land-Use and Transport for Increasing Urban Sustainability. Final Report,* 2nd ed. WSP Group, London, 2004. www.wspgroup.fi/lt/Propolis/PROPOLIS_Abstract_ Summary.pdf.
8. Shiftan, Y., and J. Suhrbier. The Analysis of Travel and Emission Impacts of Travel Demand Management Strategies Using Activity-Based Models. *Transportation,* Vol. 29, 2002, pp. 145–168.
9. Johnston, R. A., and T. de la Barra. Comprehensive Regional Modeling for Long-Range Planning: Linking Integrated Urban Models and Geographic Information Systems. *Transportation Research A,* Vol. 34, 2000, pp. 125–136.
10. Kitamura, R., S. Fujii, A. Kikuchi, and T. Yamamoto. Microsimulator Study of Daily Travel and Network Dynamics to Evaluate TDM Effectiveness for CO_2 Emissions Reduction. Presented at 77th Annual Meeting of the Transportation Research Board, Washington, D.C., 1998.
11. Miller, E. J., and M. J. Roorda. Prototype Model of Household Activity Scheduling. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1831,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 114–121.
12. Roorda, M. J., E. J. Miller, and K. M. N. Habib. Validation of TASHA: 24-Hour Activity Scheduling Microsimulation Model. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.
13. Balmer, M. *Travel Demand Modeling for Multi-agent Transport Simulations: Algorithms and Systems.* PhD dissertation. Eidgenossische Technische Hochschule, Zurich, Switzerland, 2007.
14. Charypar, D., K. W. Axhausen, and K. Nagel. Event-Driven Queue-Based Traffic Flow Microsimulation. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2003,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 35–40.
15. Gao, W. *Comparisons Between MATSim and EMME/2 on the Greater Toronto and Hamilton Area Network.* MS thesis. Department of Civil Engineering, University of Toronto, Ontario, Canada, 2009.
16. *User's Guide to Mobile 6.1 and Mobile 6.2: Mobile Source Emission Factor Model.* EPA420-R-03-010. Assessment and Standards Division, Office of Transportation and Air Quality, U.S. Environmental Protection Agency, Washington, D.C., 2003.
17. *Emissions of Air Toxics from On-Highway Sources in Canada: Estimated Impacts of Various Vehicle and Fuel Control Strategies.* Pollution Data Branch, Environment Canada, Ottawa, Quebec, Canada, 2002.
18. National Climate Data and Information Archive. *Climate Data Online (2004).* Environment Canada, Fredericton, New Brunswick, Canada, 2006. www.climate.weatheroffice.ec.gc.ca/climateData/canada_e.html.
19. Eastern Research Group, Inc. *Evaluation of Ontario Drive Clean Program.* ERG 3495.00.001.001. Ontario Ministry of the Environment, Toronto, Canada, 2005.
20. *Ontario's Drive Clean: A Summary of the First Three Years of Light-Duty Vehicle Data (1999–2001).* PIBS 4279e. Drive Clean, Toronto, Ontario, Canada, 2002.

The Travel Analysis Methods Section peer-reviewed this paper.

Integrated Bilevel Model to Explore Interaction Between Land Use Allocation and Transportation

Liyuan Zhao and Zhong-Ren Peng

The purpose of this paper is to develop a bilevel integrated dynamic model—a combination of an upper land use allocation model and a lower transportation model—to quantify the interaction between different land use allocation strategies and the transportation system. To manage the dynamic land use change in spatial and temporal dimensions, the upper-level model uses cellular automata to capture the spatial attributes of land use change, whereas the bid–rent agent model focuses on household location choice behavior. The cell-based land allocation strategy and residential location choice generated in the upper-level model are fed into the lower-level model to reflect new transportation demand, travel cost, and transportation accessibility. Then, the travel cost and transportation accessibility produced in the lower-level model are fed back into the upper-level model. To optimize land use allocation strategy, a combination of a genetic algorithm and a Frank–Wolfe algorithm is used to minimize transportation system costs. Numeric analysis of a fictitious urban area showed that the optimal land allocation with the bilevel model significantly enhanced transportation efficiency and reduced the system cost of transportation by 30.8% to 90.2%.

With rapid growth of population, land use, and transportation demand, modeling the interactions between land use and transportation is gaining attention. Relying solely on the expansion of transportation supplies has proven to be unrealistic in meeting growing transportation demand and may result in nonoptimized land development, increasing vehicle emissions and travel cost (*1*). Alternatively, adjusting the spatial distribution of future land use is a potential solution and has been reported to increase transportation efficiency, improve air quality, and control urban sprawl (*2, 3*). An optimized land use allocation strategy, therefore, will be able to match the increasing transportation demands, minimize the system cost of transportation, and balance land use and transportation. The objective of this paper is to explore the interactions between land use allocation and transportation with a proposed bilevel model.

L. Zhao, Department of Traffic and Transportation, Southwest Jiaotong University, Chengdu 610031, China. Alternate affiliation for L. Zhao: Department of Urban and Regional Planning, College of Design, Construction and Planning, University of Florida, 431 Arch Building, P.O. Box 115706, Gainesville, FL 32611-5706. Z.-R. Peng, Department of Urban and Regional Planning, College of Design, Construction and Planning, University of Florida, 431 Arch Building, P.O. Box 115706, Gainesville, FL 32611-5706. Alternate affiliation for Z.-R. Peng: School of Transportation Engineering, Tongji University, Shanghai 200092, China. Corresponding author: L. Zhao, liyuanzhao@ufl.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 14–25.
DOI: 10.3141/2176-02

Investigating the interactions between land use and transportation helps in facilitating governmental policy making and long-term planning. Change in land use over time is complicated, as it is related to changing demand and development supply, accessibility, spatial suitability of each piece of land, and decision-making results of different decision makers such as households, employers, governments, land owners, and developers. Furthermore, these changes are dynamic in spatial and temporal dimensions. Therefore, the interactions address the root causes of transportation problems and can be used to optimize the spatial structure of land use. It has been widely recognized that different purposes of land use types result in different travel demands, which directly relate to trip generation in the traditional four-step demand model; conversely, changes in transportation accessibility and travel cost affect the location choices for land development (*4, 5*).

Since the 1960s, various integrated models have investigated the interactions between land use and transportation (*6, 7*). In the literature, the integrated models are classified into eight types: spatial interaction model, spatial input–output model, mathematical optimization model, microsimulation model, discrete choice model, cellular automata (CA) model, multiagent model, and bid–rent model (*6, 7*). Compared with a single model, the integrated model is better able to represent the interactions in temporal and spatial dimensions (*8*). From the temporal perspective, with population increases, vacant lands will be developed and transportation networks will be expanded if the current land cannot meet increased transportation demands. From a spatial perspective, vacant lands in different locations are subject to different possibilities of development. In this paper, to investigate the effects of different land use allocation strategies on transportation systems, CA is used to capture the spatial drivers of land use change, whereas a bid–rent-based household agent model is deployed to analyze residential choice behavior.

CA is a spatial dynamic model used for simulating complex systems where local interactions result in global patterns (*9*). To capture the spatial and temporal characteristics of a complex urban process, CA has proven especially successful in modeling and analyzing spatial suitability in land use change, where discrete landforms occupy various spaces over time (*10, 11*). Compared with models based on transportation analysis zones (TAZs), CA can better describe land use changes. The discrete cells in CA represent change more accurately and facilitate its integration with raster-based geospatial data sets and spatial analysis in geographic information systems (GISs) (*9*).

Household location choice behavior is another major contributor to land use change. Multinomial logit models have been widely deployed to estimate the household location choice (*12, 13*); however, a single multinomial logit model inadequately represents competition

in the land use market and different economic characteristics of households (*14*). In contrast, in bid–rent models, the urban land market is controlled by an auction mechanism and the real estate or land parcel is assigned to the highest bidder. Therefore, a bid–rent-based model is used here to simulate the households' location choice in land use change.

LITERATURE REVIEW OF CELLULAR AUTOMATA AND BID–RENT MODELS

CA has been widely used to simulate complicated spatial–temporal processes (*15*). In the 1940s, Ulam suggested that the notion of a self-replicating machine would be amenable to rigorous treatment if it could be described in a "cell space" format—a geometric grid or tessellation, regular in all dimensions (*16*). Since then, CA has been used to investigate the logical nature of self-reproducible systems (*15*) and extensive experiments (*17*), to simulate phenomena in the geospatial domain, and to simulate dynamically the land use change (*9, 18, 19*).

CA well represents dynamic behaviors in temporal and spatial dimensions; however, it is limited to reflect socioeconomic characteristics and decision-making processes. CA accurately simulates changes in the state of an individual land cell, but it cannot adequately represent human decision making (*6*). Torrens believed that to be an inherent drawback of CA-motivated agent-based models, which otherwise flexibly represent individual decision-making entities' behavior and interactions (*20*). Integrating the CA and agent-based model is expected to capture the spatial drivers, human behaviors, and socioeconomic characteristics of land use change.

The bid–rent theory is a geographic economic theory that refers to how the price and demand of real estate change with distance, as they increase toward some point in the market, usually the central business district (CBD) (*21, 22*). Because travel costs rise with distance from the market (typically the CBD), rents of real estate tend to fall correspondingly. Retail establishments wish to maximize their profitability, so they are more willing to pay higher rents for land close to the CBD (*23*).

The bid–rent model has become increasingly popular among the integrated land use and transportation models (*24, 25*). The MUSSA model developed by Martínez—a land use equilibrium model—focuses on the bid choice of the competitive urban land market (*24*). Combined with a four-stage transportation model, the interactions between land use and transportation can be captured in a static equilibrium. Briceño et al. proposed a global system equilibrium model that integrates land use bidding, land use supply equilibrium, and Markovian traffic equilibrium in a hypernetwork (*26*). The integrated equilibrium model ensures the convergence of solutions under certain conditions; however, it is based on static rather than dynamic land use change. To capture the relationship between transportation and residential location, Chang and Mackett proposed a bilevel model that explores the bid–rent network equilibrium by accounting for the decision-making process of households, which is similar to an *n*-player noncooperative game following the Nash equilibrium (*7*).

In land use and transportation models based on the bid–rent theory, the changes in travel cost and transportation accessibility result in the households' relocation choice through a bidding location process, which affects the transportation demand in transportation networks (*27*). Like Briceño's integrated model (*26*) and Martínez and Henríquez's bidding model (*28*), the bid–rent model in this paper is integrated with a dynamic land use change-based CA model to explore the effect of land use allocation on transportation.

The objective of this study is to present the methodology and application of a proposed bilevel model, a combination of an upper land use allocation model and a lower transportation model. To explore the interactions between land use allocation and transportation, the upper-level model captures the process of dynamic land use change using CA and a bid–rent household agent model, whereas the lower-level model applies the user equilibrium to the transportation network. This study focused on a case study of a fictitious urban area using the proposed bilevel model. To optimize land use allocation, the system cost of transportation is minimized, using a combination of a genetic algorithm (GA) and a Frank–Wolfe algorithm.

NOTATION

b_h = monetary disutility bid for household agent of type h, according to its income;

B_{hi} = willingness-to-pay function of household agent of type h, if cell i is available for residential location;

k = land use types related to transportation, $k = 1, 2, 3, 4$ denote residential land, industrial land, commercial and services land, and institutional land (or education), respectively;

N_k = total land demand of land type k;

N_1 = residential land demand, measured by the number of households with different land type h, $N_1 = \Sigma_h N_1^h$;

p = trip purposes: the three main trip purposes ($p = 1, 2, 3$) denote home-based work trips, home-based shopping trips, and home-based school trips, respectively;

$\mathrm{Pr}_{h/i}$ = bid probability that household type h is the highest bidder for residential location i;

Pr_{hi}^{jp} = probability of a household agent of type h in residential cell i choosing cell j as a destination with purpose p;

$\mathrm{Pr}_{i/h}$ = choice probability that an alternative residential location $i \in \Omega_1$ yields the highest utility to household agent of type h;

Pr_{ik} = land development probability in cell i, from current vacant lands to a new land use type k;

r, s = origin–destination (O-D) pair in transportation network G, with origination in TAZ r and destination in TAZ s (study area is assumed to contain square cells for land use allocation; each TAZ covers a number of cells for transportation analysis; and the centroid of each TAZ is regarded as the origination or destination of an O-D pair rs);

r_i = rent of cell i;

S_{ik} = land supply of cell i for land type k; if $x_{ik} = 0$, $S_{ik} = 0$;

x_{ik} = 1, if cell i is developed into land type k; otherwise, $x_{ik} = 0$;

Ω = set of total cells in a study area, including the subset of developed cells $\Omega_1(T)$ and vacant cells $\Omega_0(T)$; both $\Omega_1(T)$ and $\Omega_0(T)$ are updated each year T;

δ_{kp} = 1, if land use type k can supply trips with purpose p; otherwise $\delta_{kp} = 0$ (main trip generation types with each land use type k are residential, origination, and destination of home-based trips; industrial, home-based work trip; commercial and services, home-based work trip and home-based shopping trip; and institutional, home-based school trip); and

λ_{ir} = characteristic variable of cell i to TAZ; if cell $i \in r$, then $\lambda_{ir} = 1$; otherwise, $\lambda_{ir} = 0$.

LOGIT-BASED CELLULAR AUTOMATA MODEL AND BID–RENT AGENT MODEL

Land Use Categories Based on Trip Generation

On the basis of the contribution to trip generation, Meyer and Miller suggested five land use types: residential, industrial, commercial and services, institutional (education), and transportation (*29*). Transportation demand is evaluated here. Transportation network changes are assumed to be unvarying. Therefore, vacant land (e.g., open land) can be developed into four trip-generation-related types denoted by k ($k = 1, 2, 3, 4$): residential, industrial, commercial and services, and institution (or education). If the current land supply cannot meet the increasing land demand, the vacant land will be developed into a trip-generation-related land type k.

Logit-Based Cellular Automata Land Use Model

In the CA model, the study area is divided into multiple cells of identical size. Each cell has a state representing the land use type. At each time increment, the cell state is stipulated by various spatial properties and updated simultaneously based on the transition rules. The spatial properties include three types: the physical attributes of the land cell, including soil quality, slope, and elevation value; the number of neighborhoods that have been developed; and the local spatial attributes, including transportation accessibility and distance to the CBD, shopping centers, education institutes, and other major public facilities.

In a traditional CA land use model, only two cell states (developed or undeveloped) are used to describe the state of land cells; however, four types of trip-generation-related land uses ($k = 1, 2, 3, 4$) are applied in this paper to relate to transportation. The mean utility function for land use change u_{ik} is an attribution from the current vacant land to each land type ($k = 1, 2, 3, 4$) in cell i. Using the multinomial logit model, u_{ik} is described as follows:

$$u_{ik} = w_{k1}\upsilon_{ik} + w_{k2}\eta_{ik} + w_{k3}\tau_{ik} \tag{1}$$

where

- w_{k1}, w_{k2}, and w_{k3} = corresponding coefficients,
- υ_{ik} = physical attributes of land cell i,
- η_{ik} = number of cells that have been developed into land type k in Moore neighborhoods of cell i, and
- τ_{ik} = transportation accessibility produced by the lower transportation model.

The mean utility function u_{ik} is associated with a stochastic error ϵ_k and is assumed to follow Gumbel infinitely divisible distribution with dispersion parameter β_k. The land development probability Pr_{ik} in cell i from current vacant land developed to a new land use type k is represented as follows:

$$\mathrm{Pr}_{ik} = \frac{\exp(\beta_k \cdot u_{ik})}{\sum_k \exp(\beta_k \cdot u_{ik})} \tag{2}$$

The higher the value of Pr_{ik} a vacant cell i has, the more possible it is to be developed into type k land. The upper land use allocation model could be used to maximize the land development probability Pr_{ik}.

Bid–Rent Model

In bid–rent theory, the urban land market is assumed to follow an auction mechanism in which the land will be developed by the highest bidder. In the bid–rent-based households' residential location choice model, the bidders (i.e., the households) are categorized into different types ($h = 1, 2, \ldots,$ and $\bar{h}$), according to socioeconomic characteristics. The bids are represented by the consumer's willingness-to-pay function for an available residential location, which is related to household income, the spatial attributes of a location, and potential transportation effects.

For households of type h and cell i available for residential location, the willingness-to-pay function B_{hi} is postulated as follows (*26*):

$$B_{hi} = -b_h + z_{hi}(\tau_{i1}) - \sum_p M_h^p \varphi_{hi}^p(t) \tag{3}$$

where

- h = total household categories,
- b_h = monetary disutility bid for household agent of type h and is proportional to the household's income,
- z_{hi} = function of τ_{i1} that captures how a household of type h values the spatial attributes of cell i,
- τ_{i1} = transportation accessibility of residential cell i,
- M_h^p = number of trips with purpose p for a household of type h, and
- $\varphi_{hi}^p(t)$ = total cost to reach purpose p, when the household of type h chooses cell i as the residential location, obtained from the logit-based trip distribution model.

The last term indicates the total transportation utility under different trip purposes p, which chooses cell i as a residential location.

The bid function $\hat{B}_{hi}$ is assumed to be a random variable. It accounts for the behavior produced by idiosyncratic differences among consumers within a cluster (*28*). The bid function can be represented by $\hat{B}_{hi} = B_{hi} + \epsilon_{hi}$, where the random item ϵ_{hi} is assumed to follow independent identically distributed (IID) Gumbel distribution with dispersion parameter θ. The bid probability, $\mathrm{Pr}_{h/i}$, the probability that the household type h is the highest bidder for location i, is given as follows:

$$\mathrm{Pr}_{h/i} = \mathrm{Pr}(B_{hi} \geq B_{h'i}) = \frac{\exp(\theta B_{hi})}{\sum_{h'} \exp(\theta B_{h'i})} \qquad \forall h = 1, 2, \ldots, \bar{h} \tag{4}$$

On the supply side, a residential location is assumed to be offered to the household with the highest payment. As a result, the rent of location r_i is determined by the expected highest bid and could be given as follows:

$$r_i = E\left[\max_h \tilde{B}_{hi}(i)\right] = \frac{1}{\theta}\ln\left(\sum_h \exp(\theta B_{hi})\right) + \frac{\gamma}{\theta} \tag{5}$$

where E = expected highest bid function and γ is a constant.

The household agents' optimal choice for a residential location is supposed to maximize the surplus between bidding price and rent, which results in the following problem: $\mathrm{Max}_{\forall i}(B_{hi} - r_{hi})$ where $\{i \mid i \in \Omega_0, x_{i1} = 1\}$ denotes the available residential cells. The rent is taken as a deterministic variable. The choice probability $\mathrm{Pr}_{i/h}$, the probability that an alternative residential location $i \in \Omega_1$ yields the highest utility to a household agent of type h, is given by:

$$\mathrm{Pr}_{i/h} = \frac{\exp(\theta(B_{hi} - r_i))}{\sum_{i \in \Omega_0} x_{i1} \exp(\theta(B_{hi} - r_i))} \tag{6}$$

$\mathrm{Pr}_{i/h}$ is the probability that a household of type h chooses a residential location i when location i is developed for residential land use.

PROPOSED BILEVEL MODEL

The proposed bilevel model (Figure 1) explores interactions between different land use allocation strategies and transportation. Figure 1 shows that the bilevel model consists of a land use allocation model (upper level) and a transportation model (lower level). With a GA, the upper-level model produces a land use allocation strategy and a new number of households in the residential location. The new travel demand between zones is updated from results of the upper-level model and is fed into the lower transportation model, a classic user equilibrium transportation model. With the Frank–Wolfe algorithm, travel cost and transportation accessibility are produced in the lower-level model and fed into the upper-level model. The resulting land use allocation suitability will generate more accurate land use allocations. Thus, an optimal land use allocation strategy with minimum transportation system cost is created.

At the cell scale, the bilevel model integrates residential choice behavior into the logit-based CA model. In the upper-level model, the vacant lands with higher total land development spatial suitability will first be generated into trip-related land types. On the basis of households' residential location choice, the vacant land cells with a higher willingness-to-pay value and lower rent are preferred and will be the first to be developed into residential land for the newly increased number of households.

Upper Land Use Model

The upper land use model can be formulated as follows:

$$ZU = \min_{(x_{ik}, N_{ik}, b, r)} \begin{pmatrix} -\sum_{i\in\Omega_0}\sum_{k} x_{ik} N_{ik} \mathrm{Pr}_{ik}(t) + \sum_{h} N_1^h b_h + \sum_{i\in\Omega_0} x_{i1} S_{i1} r_i \\ +\frac{1}{\theta}\sum_{h}\sum_{i\in\Omega_0} x_{i1} \exp(\theta(B_{hi}(b,t) - r_i)) \end{pmatrix} \tag{7}$$

such that

$$x_{ik} = \begin{cases} 0 \\ 1 \end{cases}$$

and

$$\sum_{k} x_{ik} \le 1 \tag{8}$$

$$\sum_{i} N_{ik} = N_k \qquad \forall k = 1, 2, 3, \text{and } 4 \tag{9}$$

$$N_{ik} \le S_{ik} \qquad \forall k, 1 \tag{10}$$

where ZU is the objective function in the upper land use model and S_{i1} denotes the maximum bounded residential land supply in vacant cell i. The first term of Equation 7 aims to maximize the total land development spatial suitability and represents the allocation of a new land type k ($k = 1, 2, 3, 4$) into the vacant cell i ($i = 1, 2, \ldots, \bar{i}$). The characteristic variable x_{ik} indicates whether the vacant cell i will

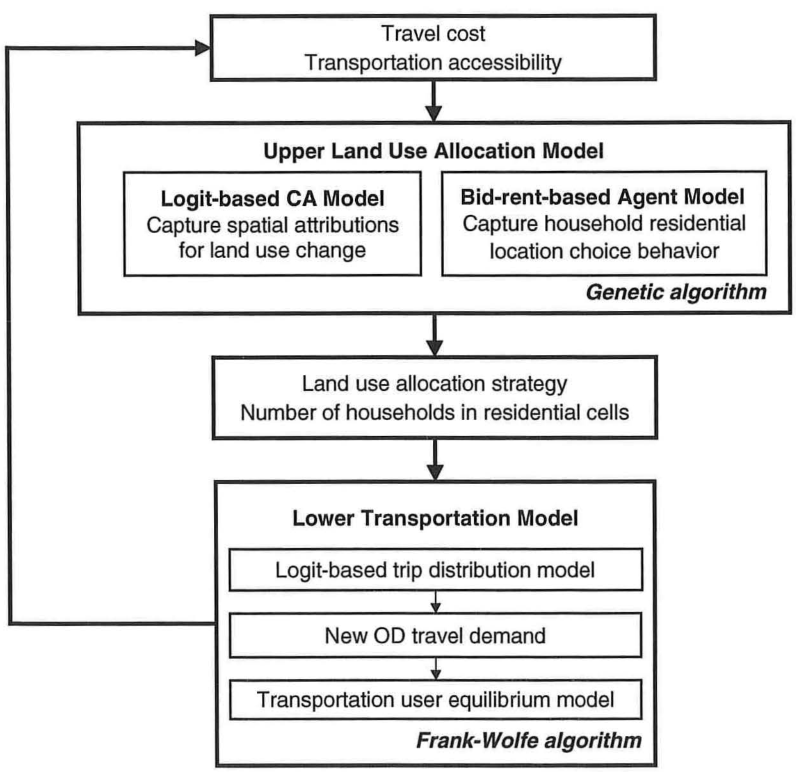


FIGURE 1 Overview of proposed bilevel model, a combination of upper land-use allocation model and lower transportation model.

be developed into land type k. When cell i is developed into residential land, $x_{i1} = 1$. N_{ik} is the allocated amount of land type k in cell i. $\Pr_{ik}(t)$ is relevant to travel cost t and denotes the suitability of vacant cell i to be developed into land type k.

The bid-rent-based household location choice model is expressed in the last three terms of Equation 7. The second term represents the total bidding utility of different household agents. The third term is the utility of supply location, depending on rents. The last term is used to ensure equivalence with the bid–rent model.

Equation 8 indicates that each land cell can be developed into only one land use type. Equation 9 maintains the balance between land use demand and allocation. Equation 10 ensures that the allocated number of land cells falls within the maximum bounded land supply S_{ik}.

In the upper land use model, the equivalence with the bid–rent model is deduced as follows:

First, from

$$\frac{\partial ZU}{\partial b_h} = 0$$

is the number of type h household agents:

$$N_1^h = \sum_i x_{i1} \exp\left(\theta\left(B_{hi}\left(b_h, t\right) - r_i\right)\right) \tag{11}$$

When cell i is developed into residential land ($k = 1$), defining the number of household agents of type h, choosing cell i as residential:

$$N_{i1}^h = x_{i1} \exp\left(\theta\left(B_{hi}\left(b_h, t\right) - r_i\right)\right) \tag{12}$$

Substituting Equation 12 into Equation 11 results in

$$N_1^h = \sum_i N_{i1}^h \tag{13}$$

Combining Equations 11, 12, and 13 results in

$$N_{i1}^h = N_1^h \frac{x_{i1} \exp\left(\theta\left(B_{hi}\left(b_{hi}, t\right) - r_i\right)\right)}{\sum_i x_{i1} \exp\left(\theta\left(B_{hi}\left(b_h, t\right) - r_i\right)\right)} \tag{14}$$

Combined with the results of residential land allocation, Equation 14 yields the choice probability $\Pr_{i/h}$ in the stochastic bid–rent location choice model (Equation 6).

Second, with

$$\frac{\partial ZU}{\partial r_i} = 0$$

and Equation 12, the following is created:

$$x_{i1} S_{i1} = \sum_h x_{i1} \exp\left(\theta\left(B_{hi}\left(b_h, t\right) - r_i\right)\right) = \sum_h N_{i1}^h \tag{15}$$

With a combination of Equations 12 and 15, the following is created:

$$N_{i1}^h = \frac{x_{i1} S_{i1} \cdot \exp\left(\theta B_{hi}\right)}{\sum_h \exp\left(\theta B_{hi}\right)} \tag{16}$$

Therefore, the bid probability $\Pr_{h/i}$ in Equation 4 is updated by the new N_{i1}^h in Equation 16. When location i is developed into residential land, the households with the highest bidder for location i will choose location i.

Lower Transportation Model

The lower-level model in the proposed bilevel model consists of a logit-based trip distribution model and transportation user equilibrium model. New O-D travel demand produced from the logit-based trip distribution model is then fed into the transportation user equilibrium model.

Logit-Based Trip Distribution Model

To link the upper land use model and lower transportation model, the logit-based trip distribution model is used to produce travel demand. The travel demand q_{rs} for the O-D pair rs (TAZ to TAZ) contains existing travel demand $\bar{q}_{rs}$ and extra travel demand $\hat{q}_{rs}$, which is generated from the new households in the upper land use model. The travel demand q_{rs} is formulated as follows:

$$q_{rs} = \bar{q}_{rs} + \tilde{q}_{rs} \tag{17}$$

For all cells in study area Ω, the set of cells that can be developed into the destinations of trips with purpose p, is denoted as $\Omega_p = \{j | x_{jk}\ \delta_{kp} = 1, j \in \Omega\}$. For households of type h in cell i, the utility $A_{hi}^{jp}(t)$ of choosing cell j ($j \in \Omega_p$) as the destination of a trip with purpose p consists of two parts: (*a*) t_{rs} ($i \in r, j \in s$), the travel cost from cell i to j, generated from O-D pair travel cost based on a zone-cell matrix, and (*b*) ψ_{hjp}, the attractiveness of cell j.

$$A_{hi}^{jp}\left(t\right) = -t_{rs} + \psi_{hjp} \qquad i \in r, j \in s \tag{18}$$

It assumes that the utility $A_{hi}^{jp}(t)$ follows IID Gumbel distribution with dispersion parameter ϕ. A_{hi}^{jp} is the probability that a household of type h in residential cell i chooses cell j as a destination with trip purpose p. $\Pr_{hi}^{jp}$ is formulated as follows:

$$\Pr_{hi}^{jp} = \frac{\exp\left(\phi \cdot A_{hi}^{jp}\left(t\right)\right)}{\sum_{j \in \Omega_p} \sum_k \exp\left(\phi \cdot A_{hi}^{jp}\left(t\right)\right)} \tag{19}$$

In Equation 3, $\varphi_{hi}^{p}(t)$ is the total cost to reach trip purpose p, when household type h selects cell i as the residential location. The total cost $\varphi_{hi}^{p}(t)$ can be represented as follows:

$$\varphi_{hi}^{p}\left(t\right) = \sum_j \Pr_{hi}^{jp} \cdot A_{hi}^{jp}\left(t\right) \tag{20}$$

The total travel demand $\hat{q}_{ij}$ from new residential cell i to destination cell j is denoted as follows:

$$\tilde{q}_{ij} = \sum_h \sum_p N_{i1}^h M_h^p \Pr_{hi}^{jp} \qquad i \in \Omega_0, j \in \Omega \tag{21}$$

where $N_{i1}^h M_h^p$ is the number of trips with purpose p for household type h in residential cell i.

From the relationship between zones and cells, the zone-based extra travel demand $\hat{q}_{rs}$ is represented as follows:

$$\tilde{q}_{rs} = \sum_{i \in \Omega_0} \sum_{j \in \Omega} \lambda_{ir} \lambda_{js} \tilde{q}_{ij} \tag{22}$$

With Equation 17, in the lower-level model, the O-D demand q_{rs} is updated.

Transportation User Equilibrium Model

In the transportation user equilibrium model, the travel demand q_{rs} is updated using the land use allocation strategy and households' residential location choice, which are generated from the upper land use model (Figure 1). The Frank–Wolfe algorithm (*30*) is used to solve the user equilibrium problem (Equation 23) in this paper. Traditionally, all used routes in a transportation network share the same travel costs for the same O-D pair. The user equilibrium model in this paper is simplified as the following programming problem (*31*):

$$\min Z = \sum_a \int_0^{f_a} t_a(v)dv \tag{23}$$

such that

$$\sum_l f_l^{rs} = q_{rs} \qquad \forall r, s \tag{24}$$

$$f_l^{rs} \geq 0 \qquad \forall r, s, l \tag{25}$$

$$f_a = \sum_{rs}\sum_l f_l^{rs}\delta_{a,1}^{rs} \qquad a \in A \tag{26}$$

where a is the link in a transportation network and A is the set of all studied links. f_a is the flow on link a. $\delta_{a,l}^{rs}$ equals 1, if link a belongs to path l, which is between the O-D pair rs. $t_a(v)$ denotes the link travel time when it carries flow f_a. The Bureau of Public Roads function (*32*) is deployed to describe the link travel time function:

$$t_a = t_a^0\left[1 + 0.15\left(\frac{v_a}{C_a}\right)^4\right] \tag{27}$$

where t_a^0 is free-flow travel time, C_a is link capacity, and v_a is actual volume on link a.

Model Structure of Bilevel Model

To optimize the interactions between land use allocation and transportation, a self-coded MATLAB program was developed based on the bilevel model (Figure 2). A GA was deployed to solve the upper land use allocation model and the Frank–Wolfe algorithm was used in the lower transportation model. GA aims to maximize the combined spatial land development suitability in the CA model and households' residential choice preference (*33, 34*). The Frank–Wolfe algorithm,

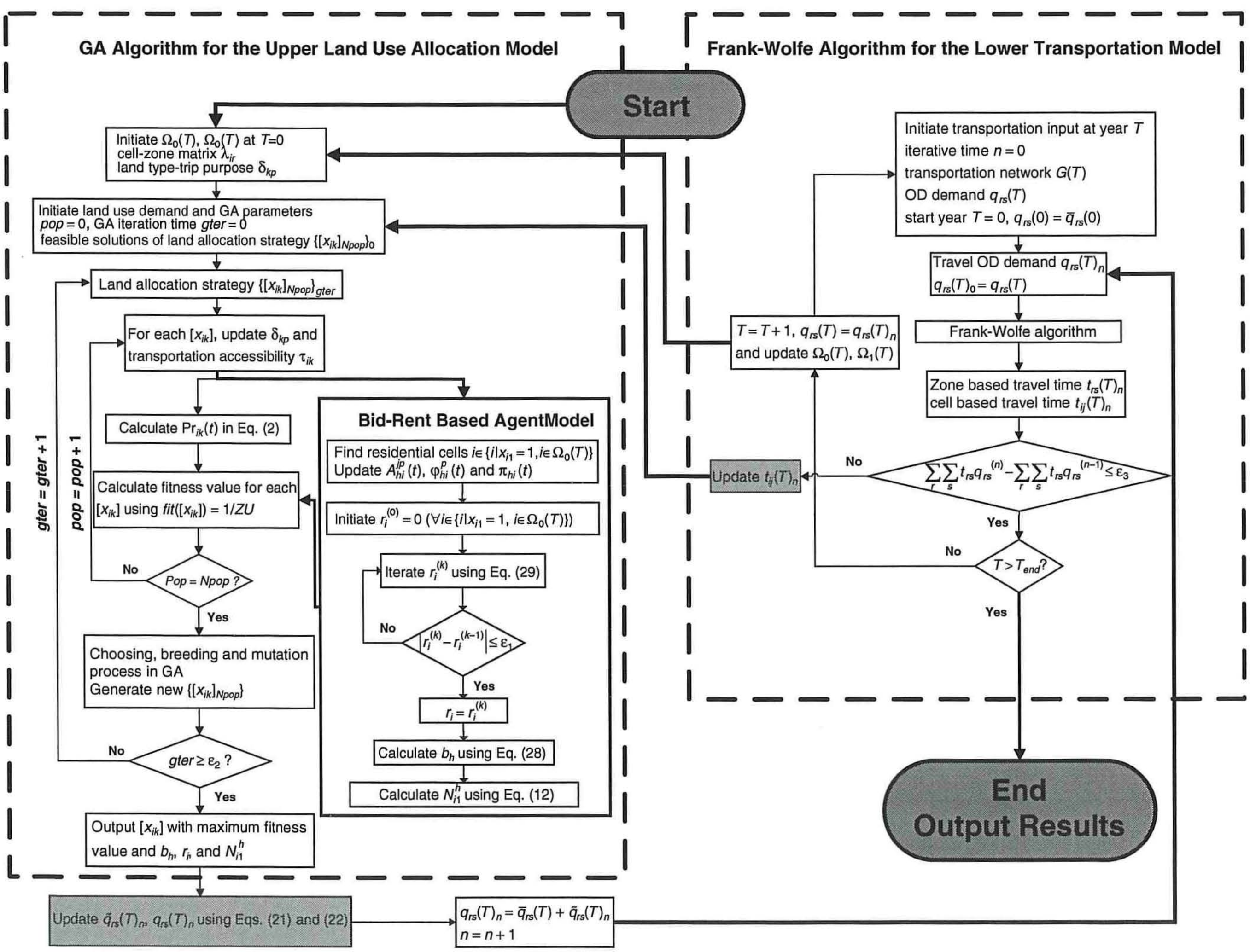


FIGURE 2 Flowchart of model structure for bilevel model.

also known as the convex combination algorithm, is the predominant algorithm used to solve traffic assignment problems in transportation planning studies (*35, 36*).

RESULTS AND DISCUSSION

In this section, a fictitious urban area with nine TAZs is examined over 3 years, using the proposed bilevel model. Figure 3 shows the initial land use for each cell in the initial year $T = 0$. To simulate the dynamic land use change using CA, each TAZ in the study area is further divided into 25 square cells of identical size. Each cell has only one state that represents one of the five land use types (vacant, residential, industrial, commercial and service, and institutional). There are 24 links connecting the centroid of each zone in the transportation network. With time increasing (from year $T = 0$ to year $T = 3$), the population and land demands in each zone will increase correspondingly.

Model Application of Bilevel Model

To simplify the bilevel model, the physical attributes of land cell υ_{ik} (Equation 1) in the fictitious study area are neglected ($w_{k1} = 0$) in this paper. Only Moore neighborhoods (eight neighboring cells) and transportation accessibility in the CA are considered ($w_{k2} = w_{k3} = 1$). Neighborhoods' attribution η_{ik} represents the number of cells developed into land type k in Moore neighborhoods, which can be generated from the land use situation each year. The transportation accessibility τ_{ik} is measured by the inverse of travel time to the nearest land cell with land type k.

To determine the parameters in the upper land use model (Equation 7), the new land demand of each type (N_k) for each year is assumed as 1,000, 1,000, 800, and 500, and the maximal land supply of each cell (S_{ik}) is assumed as 100. Therefore, the total increase in land demand is 33 land cells per year. Households are classified into four types ($h = 1, 2, 3, 4$), with 250 households for each type ($N_1^h = 250$). It is assumed that rich households generate more trips than poor ones. For simplicity, different socioeconomic characteristics for each household type are represented by the number of trips M_p^h.

The values of M_p^h for each trip generation type (work, shopping, school) are (5, 4, 3), (4, 3, 2), (3, 2, 2), and (2, 2, 1), respectively. The spatial attributes $z_{hi}(\tau_i 1)$ in the willingness-to-pay function (Equation 3) are therefore supposed to be the same as $\tau_i 1$.

In the distribution part (Equation 18), the attractiveness ψ_{hjp} is supposed to be the same for each cell; therefore, only travel cost t_{rs} is considered in Equations 18, 19, and 20. In the logit model, the parameters β_k, θ, and ϕ are assumed to be 1.0, 0.8, and 1.0, respectively. Alternatively, they can be generated in the multinomial logit regression model, if actual land use data are used.

Optimization of Land Use Allocation

With a combination of a GA and the Frank–Wolfe algorithm in the proposed bilevel model, the land use allocation is optimized by searching an optimal land use allocation strategy $[x_{ik}(T)]$ with a minimal total transportation system cost. The optimal land use allocation strategy aims to maximize the land use suitability in CA, household residential location choice behavior based on the bid–rent mechanism in land market, and user equilibrium in the transportation network. In each time step from year $T = 1$ to year $T = 3$, a new land status in year $T + 1$ was updated by the produced optimal land use allocation strategy $[x_{ik}(T)]$ in year T. In contrast, the worst land allocation with a maximal total system cost of transportation in year $T + 1$ was obtained based on a worst land allocation strategy in the previous year T.

Table 1 compares the optimal and worst land use allocation strategies $[x_{ik}(T)]$ for the 3 years studied ($T = 1, 2, 3$). The total system cost of transportation increased sharply (e.g., from 1,734.5 in year $T = 2$ to 9,166.8 in year $T = 3$ for the worst land allocation strategy) as the population and land demand increased in the same transportation network. Compared with the worst-case scenario, the optimal land allocation strategy each year reduced the system cost of transportation by 30.8%, 59.7%, and 90.2%, respectively, over the 3 years. With both population and land demand increasing from year $T = 1$ to year $T = 3$ (Table 1), the proposed bilevel model more effectively reduced travel cost and therefore better optimized land allocation. The more land demand and population, the more significantly the bilevel model

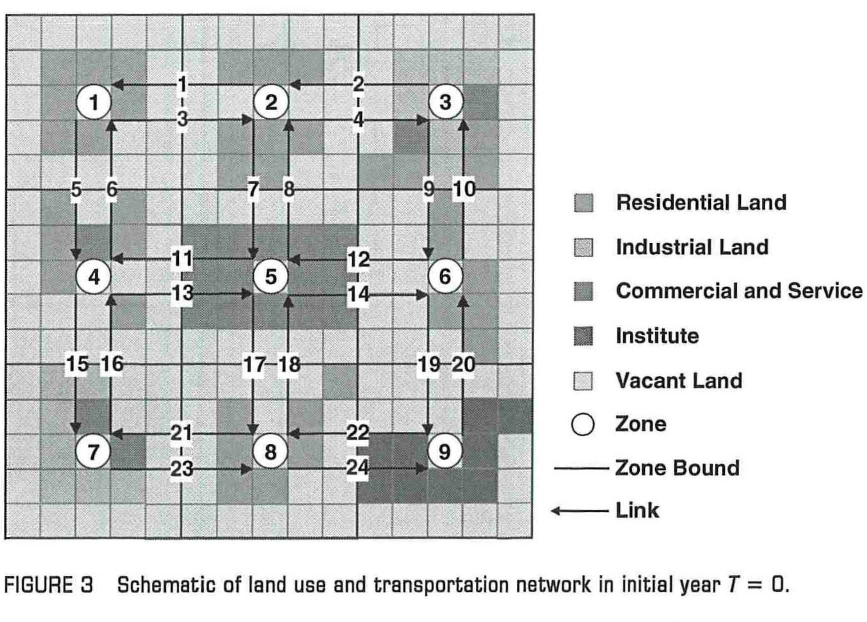


FIGURE 3 Schematic of land use and transportation network in initial year $T = 0$.

TABLE 1 Comparison of Optimal and Worst Land Use Allocation Strategies

Year	Increased Number of Household (units)	Increased Land Demand (cells)	Total System Cost of Transportation[a] Optimal Land Allocation (OL) (*1,000)	Worst Land Allocation (WL) (*1,000)	Cost Reduction[b] (%)
$T = 1$	1,000	33	482.3	696.9	30.8
$T = 2$	1,000	33	698.7	1,734.5	59.7
$T = 3$	1,000	33	900.2	9,166.8	90.2

[a]Total system cost of transportation is measured by a sum of the products between O-D demand and the corresponding travel cost in transportation networks.

[b]$\text{cost reduction} = \dfrac{(\text{system cost})_{WL} - (\text{system cost})_{OL}}{(\text{system cost})_{WL}}$

enhanced transportation efficiency. Therefore, the bilevel model is expected to perform better in a more complicated land use and transportation network (e.g., real land use data) than in the nine-zone urban area in this paper.

Figure 4 compares the optimal and worst land use allocation strategies in the 3 years studied (Table 1). In the optimal land use allocation strategy (Figure 4*a*, *c*, and *e*), new developed land cells were dispersed in the whole study area. In contrast, in the worst cases (Figure 4*b*, *d*, and *f*), almost all newly developed land cells were allocated near the existing land with the same land use type. In the optimal land allocation strategy, new land cells are more dispersed, whereas they are more agglomerate in the worst case. Extensively allocating land cells close to the area with the same land type results in a high system cost of transportation.

As numerical results show, optimizing land use allocation with the bilevel model is able to reduce transportation cost. This finding agrees with similar conclusions reported in previous studies based on integrated land use and transportation models (*7*). Unlike the other models, the bilevel model in this paper is able to provide a quantitative analysis of future land use and suggests in which directions the land use could be optimized to minimize the total system cost of transportation.

Households' Residential Location Choice Based on Bid–Rent Theory

In this section, the socioeconomic characteristics of households in residential location choice were captured by the bid–rent theory. In the bid–rent theory, the willingness-to-pay function and rent are directly related to transportation travel costs. Lower travel cost will lead to a higher bidding price for locations, and higher travel cost will lead to a lower bidding price for locations.

Figure 5*a* illustrates the optimal distribution of new residential locations for each year ($T = 1, 2, 3$). The households prefer residential locations close to the center of transportation networks (e.g., the CBD). In initial year $T = 0$, vacant lands close to the center were first developed into residential lands. In year $T = 3$, fewer vacant lands were available near the center; therefore, more residential locations were selected in the edge zones of the study transportation network (e.g., Locations 6, 7, and 8 in TAZ 3).

Figure 5 shows the distribution of different households in new optimal residential locations. In the bid–rent theory, because of their lowest bidding price, poorer households have fewer choices to select from limited residential locations. In contrast, richer households are more flexible to select residential locations. For instance, in year $T = 1$, rich households (Types 1 and 2) were evenly distributed in new residential lands (Figure 5*b*); however, the poorest households (Type 4) had limited choices, most of which were located in residential Locations 2 and 4 as indicated in Figure 5*b*.

Figure 5 also shows that most of the poorest households (Type 4) choose residential locations in the central TAZ (Locations 2 and 4 in TAZ 5). The central TAZ contained only commercial lands in year $T = 1$. So the central TAZ was inconvenient to the destinations of other trip purposes and had higher total travel costs compared with other TAZs with diverse land types. In the bid–rent theory, higher travel cost results in lower rents for residential locations. Therefore, the poorest households selected the residential locations close to the central TAZ. In contrast, the richer households preferred the residential locations in TAZ 9, where different land types were available for multiple trip purposes.

Effects of Land Use Allocation on Transportation Network

Transportation affects land use allocation; conversely, land use allocation could also affect transportation performance. In this section, the effects of land use allocation on the transportation network are examined. In general, the performance of transportation networks could be evaluated by link saturation, the ratio between assigned flow and capacity. Typically, high link saturation (>85%) is regarded as having traffic congestion (*37*).

Figure 6 compares link saturations in the optimal and worst land use allocations. With the population and land use increase from year $T = 1$ to year $T = 3$, transportation demand increased and the study transportation network started to congest. For instance, most links will be saturated (>85%) and over capacity in year $T = 3$ for both the optimal and the worst land use allocations (Figure 6).

Optimizing land use allocation helps to relieve long-term traffic congestion. High saturation (>300%) of Links 19 and 24 occurred in the case of the worst land use allocation (Figure 6*b*), whereas relatively smooth traffic was observed in the case of the optimal land use allocation (Figure 6*a*). Furthermore, the land use allocation was optimized by the proposed bilevel model; however, most links in the transportation network are close to or over the capacity (>85%) in year $T = 3$. This finding shows that relying solely on optimization by the proposed bilevel model is insufficient to solve the increasing

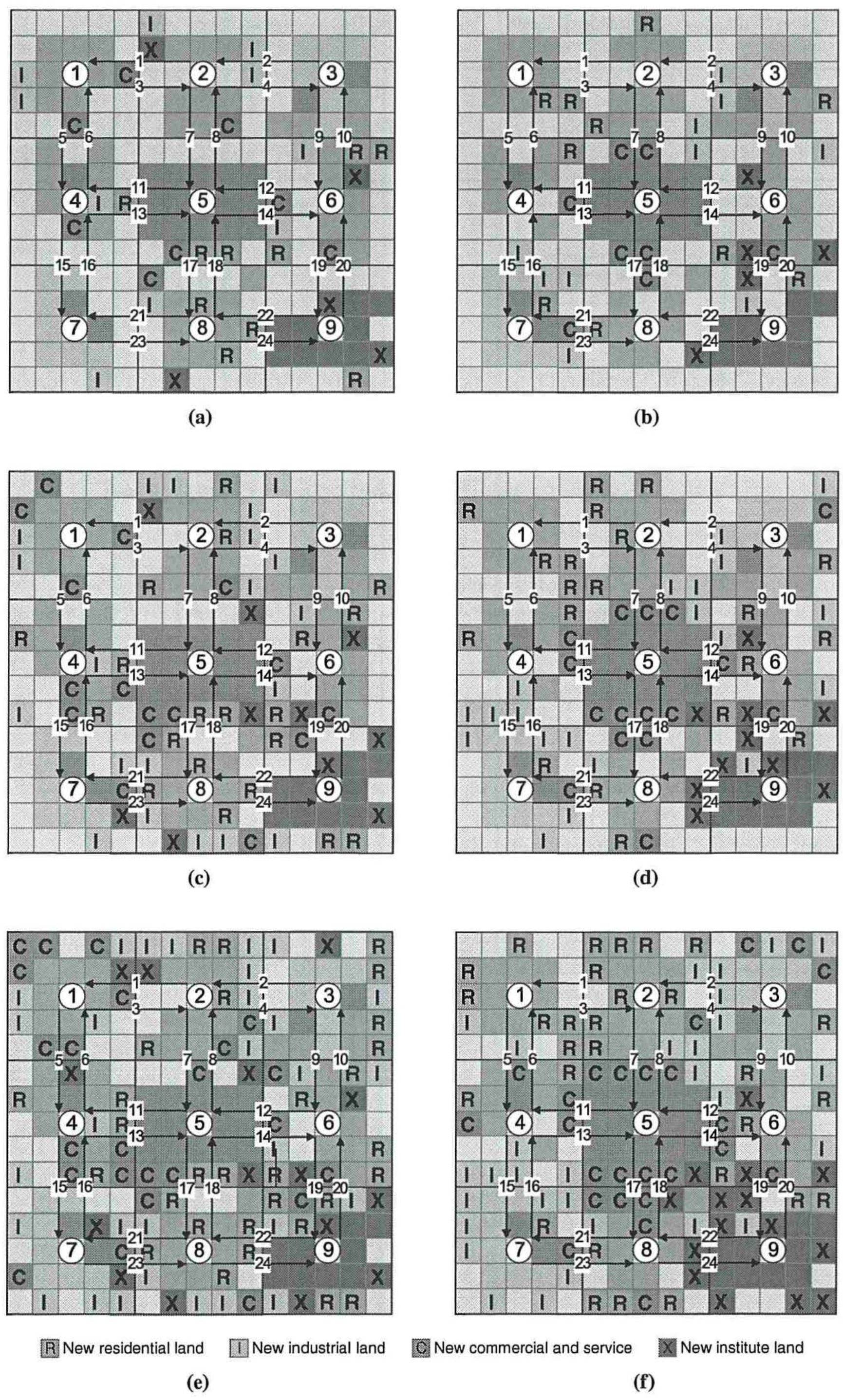


FIGURE 4 Comparison of optimal and worst land use allocation strategies for study transportation network: (*a*) optimal case in year T = 1, (*b*) worst case in year T = 1, (*c*) optimal case in year T = 2, (*d*) worst case in year T = 2, (*e*) optimal case in year T = 3, and (*f*) worst case in year T = 3.

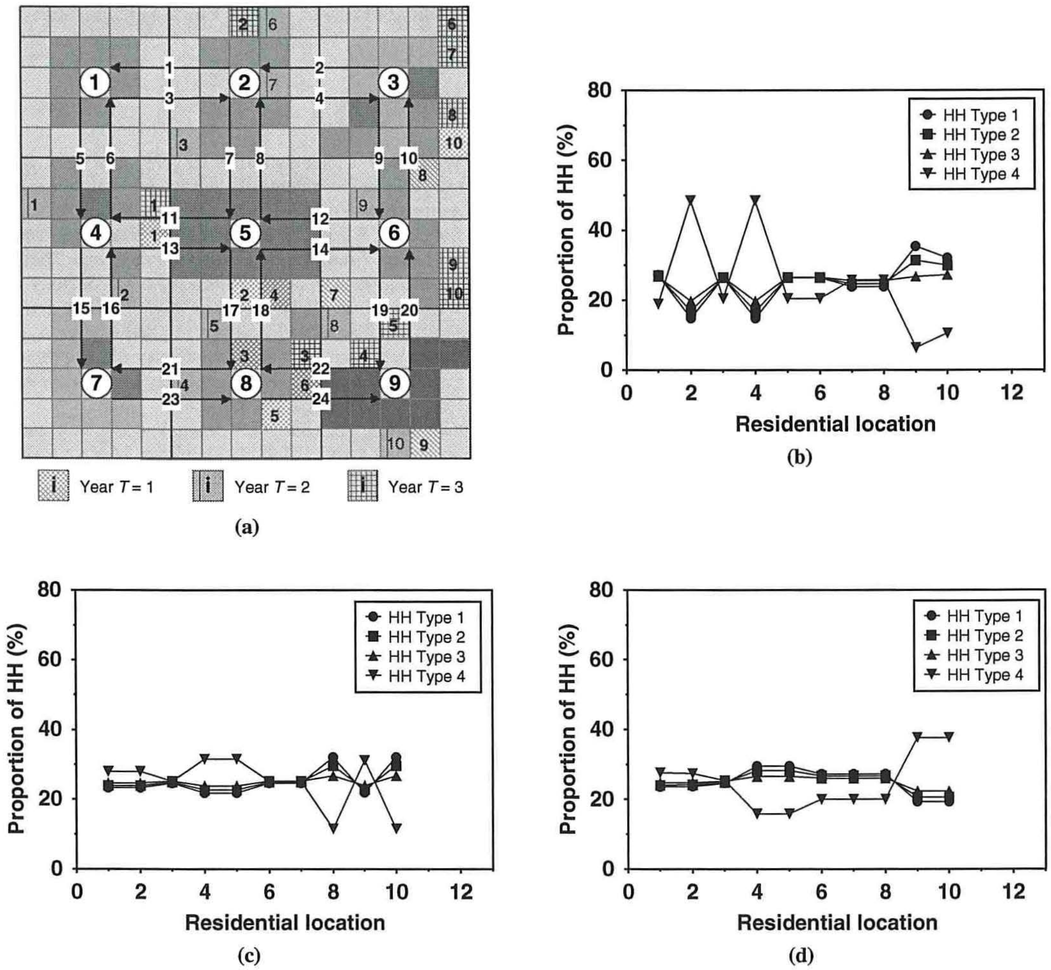


FIGURE 5 Optimal land use allocation of study transportation network: (*a*) overview of new residential land, (*b*) proportion of new households in year $T = 1$ (Household Type 1, richest; Type 2, rich; Type 3, poor; Type 4, poorest), (*c*) proportion of new households in year $T = 2$, and (*d*) proportion of new households in year $T = 3$.

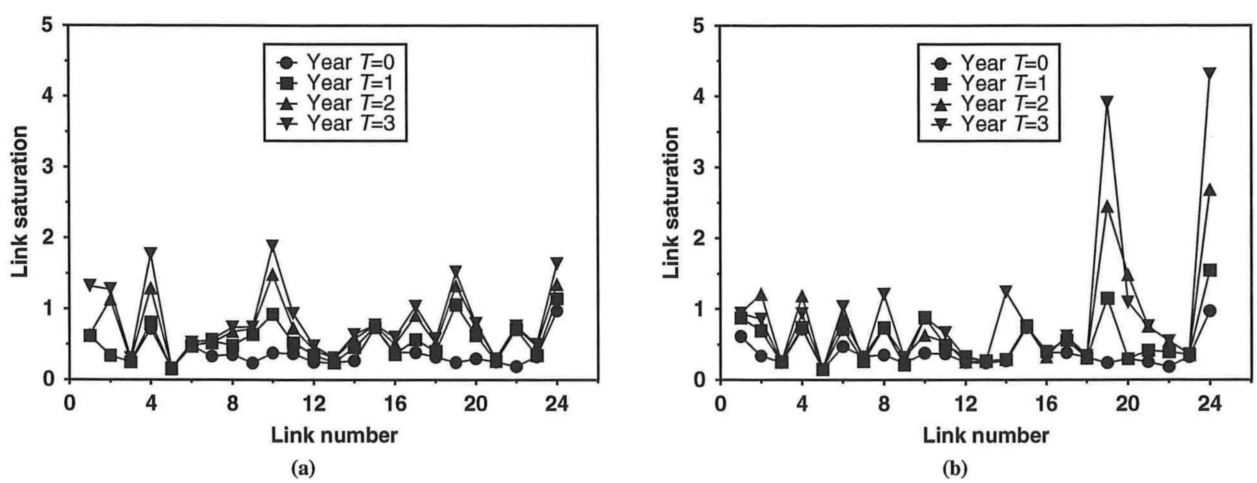


FIGURE 6 Link saturation in study transportation network: (*a*) optimal land use allocation and (*b*) worst land use allocation.

land use and transportation demand. Expanding the transportation network is an alternative solution.

Outlook for Bilevel Model

Although the proposed bilevel model showed promising results and successfully optimized land use allocation with increasing transportation demand in the fictitious urban area, there is still room for improvement. In this study, only home-based trips were considered based on household choice and three trip purposes (work, shopping, school); however, non-home-based trips and trips with other purposes (e.g., truck and taxi) also play an important role in urban transportation. Non-home-based trips typically account for 25% to 30% of travel by individuals in urban areas (*38*). An improved model could estimate trip production and attractions for all the above-mentioned trip purposes. Furthermore, a comprehensive integrated bilevel model could consider other factors, such as behavior of land developers and owners, economic development, transportation policy, and environmental effects. Implementing GIS and real land use data into the model will provide diverse and in-depth references for discussing the advantages and disadvantages of the proposed bilevel model.

CONCLUSIONS

A bilevel model is proposed to investigate the interactions between land use allocation and transportation networks. In the upper land use allocation model, CA is deployed to simulate dynamic land use change in the spatial and temporal dimensions, whereas the bid–rent theory represents different socioeconomic characteristics among households as well as competition in the land use market. The lower transportation model consists of the logit-based trip distribution model and classic user equilibrium of the transportation network model. The upper-level model generates a cell-based land allocation strategy and residential location choice, which is fed into the lower transportation model to update the zone-based travel demand. New travel cost and transportation accessibility will be produced from the lower transportation model and fed back to the upper-level model. A combination of a GA and a Frank–Wolfe algorithm is used to search an optimal land use allocation strategy with minimal system cost of transportation.

The proposed bilevel model investigated the effects of transportation on land use allocation. The optimized land allocation of a nine-zone network by the model significantly enhanced transportation efficiency and reduced the system cost of transportation by 30.8% to 90.2%. The model provides a methodology to assess land use patterns in terms of transportation efficiency and thus provides suggested land use changes for future land use planning and zoning. Furthermore, the model provides a methodology to assess and quantify the impacts of land use patterns on future transportation networks and therefore provides information for long-term land use and transportation planning.

ACKNOWLEDGMENTS

The authors thank Sarah Perch for critically reading the manuscript, making helpful comments, and improving the text. This research was supported in part by a U.S. National Science Foundation award, a National Natural Science Foundation of China award, a China National High Technology Research and Development Program 863 award, the Center for Multimodal Solutions for Congestion Mitigation at the University of Florida, and the Florida Department of Transportation.

REFERENCES

1. Waddell, P. Modeling Urban Development for Land Use, Transportation, and Environmental Planning. *Journal of the American Planning Association,* Vol. 68, No. 3, 2002, pp. 297–314.
2. Badoe, D. A., and E. J. Miller. Transportation-Land-Use Interaction: Empirical Findings in North America, and Their Implications for Modeling. *Transportation Research D,* Vol. 5, No. 4, 2000, pp. 235–263.
3. Webster, F. V., P. H. Bly, and N. J. Paulley. *Urban Land-Use and Transport Interaction.* Avebury, United Kingdom, 1994.
4. Boarnet, M., and R. Crane. The Influence of Land Use on Travel Behavior: Specification and Estimation Strategies. *Transportation Research A,* Vol. 35, No. 9, 2001, pp. 823–845.
5. Kitamura, R., E. I. Pas, C. V. Lula, T. K. Lawton, and P. E. Benson. The Sequenced Activity Mobility Simulator (SAMS): An Integrated Approach to Modeling Transportation, Land Use and Air Quality. *Transportation,* Vol. 23, No. 3, 1996, pp. 267–291.
6. Waddell, P., and G. F. Ulfarsson. Introduction to Urban Simulation: Design and Development of Operational Models. In *Handbooks in Transport, Vol. 5. Transport Geography and Spatial Systems* (P. Stopher, K. J. Button, K. E. Haynes, and D. A. Hensher, eds.), Elsevier Science, Maryland Heights, Mo., 2004, pp. 203–236.
7. Chang, J. S., and R. L. Mackett. A Bi-Level Model of the Relationship Between Transport and Residential Location. *Transportation Research B,* Vol. 40, No. 2, 2006, pp. 123–146.
8. Shaw, S. L., and X. Xin. Integrated Land Use and Transportation Interaction: A Temporal GIS Exploratory Data Analysis Approach. *Journal of Transport Geography,* Vol. 11, No. 2, 2003, pp. 103–115.
9. Bone, C., and S. Dragicevic. *Sensitivity of a Fuzzy-Constrained Cellular Automata Model of Forest Insect Infestation.* Citeseer, 1964.
10. Batty, M., Y. Xie, and Z. Sun. Modeling Urban Dynamics Through GIS-Based Cellular Automata. *Computers, Environment and Urban Systems,* Vol. 23, No. 3, 1999, pp. 205–233.
11. Li, X., and A. G. Yeh. Modelling Sustainable Urban Development by the Integration of Constrained Cellular Automata and GIS. *International Journal of Geographical Information Science,* Vol. 14, No. 2, 2000, pp. 131–152.
12. Dieleman, F. M., M. Dijst, and G. Burghouwt. Urban Form and Travel Behaviour: Micro-Level Household Attributes and Residential Context. *Urban Studies,* Vol. 39, No. 3, 2002, pp. 507–527.
13. Waddell, P., A. Borning, M. Noth, N. Freier, M. Becke, and G. Ulfarsson. Microsimulation of Urban Development and Location Choices: Design and Implementation of Urbanism. *Networks and Spatial Economics,* Vol. 3, No. 1, 2003, pp. 43–67.
14. Lipscomb, C., and M. Farmer. Household Diversity and Market Segmentation Within a Single Neighborhood. *Annals of Regional Science,* Vol. 39, No. 4, 2005, pp. 791–810.
15. White, R., and G. Engelen. Cellular Automata and Fractal Urban Form: A Cellular Modelling Approach to the Evolution of Urban Land-Use Patterns. *Environment and Planning A,* Vol. 25, 1993, pp. 1175–1193.
16. Ulam, S. Random Processes and Transformations. *Proc., 1950 International Congress on Mathematics,* No. 2, 1952, pp. 264–275.
17. Wolfram, S., and M. Gad-el-Hak. A New Kind of Science. *Applied Mechanics Reviews,* Vol. 56, 2003, p. B18.
18. Clarke, K. C., and L. J. Gaydos. Loose-Coupling a Cellular Automaton Model and GIS: Long-Term Urban Growth Prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science,* Vol. 12, No. 7, 1998, pp. 699–714.
19. Li, X., and A. G. Yeh. Neural-Network-Based Cellular Automata for Simulating Multiple Land Use Changes Using GIS. *International Journal of Geographical Information Science,* Vol. 16, No. 4, 2002, pp. 323–343.
20. Torrens, P. M. New Advances in Urban Simulation: Cellular Automata and Multi-Agent Systems as Planning Support Tools. In *Planning Support Systems in Practice* (S. S. Geertman, ed.), Springer-Verlag, London, 2002.
21. Wheaton, W. C. Bid Rent Approach to Housing Demand. *Journal of Urban Economics,* Vol. 4, No. 2, 1977, pp. 200–217.

22. Shieh, Y. N. An Early Use of Bid Rent Functions. *Urban Studies,* Vol. 40, No. 4, 2003, pp. 791–795.
23. Chang, J. S. Multiclass Bid–Rent Network Equilibrium Model. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2003,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 112–119.
24. Martinez, F. MUSSA: Land Use Model for Santiago City. In *Transportation Research Record 1552,* TRB, National Research Council, Washington, D.C., 1996, pp. 126–134.
25. Martinez, F. J. The Bid-Choice Land-Use Model: An Integrated Economic Framework. *Environment and Planning A,* Vol. 24, No. 6, 1992, pp. 871–885.
26. Briceño, L., R. Cominetti, C. E. Cortés, and F. Martínez. An Integrated Behavioral Model of Land Use and Transport System: A Hyper-Network Equilibrium Approach. *Networks and Spatial Economics,* Vol. 8, No. 2, 2008, pp. 201–224.
27. Martinez, F. Access: The Transport-Land Use Economic Link. *Transportation Research B,* Vol. 29, No. 6, 1995, pp. 457–470.
28. Martínez, F., and R. Henríquez. A Random Bidding and Supply Land Use Equilibrium Model. *Transportation Research B,* Vol. 41, No. 6, 2007, pp. 632–651.
29. Meyer, M. D., and E. J. Miller. *Urban Transportation Planning: A Decision-Oriented Approach.* McGraw-Hill Science/Engineering/Math, New York, 2001.
30. Sheffi, Y. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods.* Prentice-Hall, Englewood Cliffs, N.J., 1985.
31. Beckman, M., C. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation.* Yale University Press, New Haven, Conn., 1956.
32. Hamdouch, Y., M. Florian, D. W. Hearn, and S. Lawphongpanich. Congestion Pricing for Multi-Modal Transportation Systems. *Transportation Research B,* Vol. 41, No. 3, 2007, pp. 275–291.
33. Zhang, W. Q., S. Afshar, and T. M. Monro. A Genetic Algorithm Based Approach to Fiber Design for High Coherence and Large Bandwidth Supercontinuum Generation. *Optics Express,* Vol. 17, No. 21, 2009, pp. 19,311–19,327.
34. Jia, W., L. Y. Jiang, H. P. Li, G. G. Zheng, and X. Y. Li. Genetic Algorithm for Band Gap Optimization Under Light Line in Two-Dimensional Photonic Crystal Slab. *Optica Applicata,* Vol. 39, No. 3, 2009, pp. 481–488.
35. Ahipasaoglu, S. D., P. Sun, and M. J. Todd. Linear Convergence of a Modified Frank–Wolfe Algorithm for Computing Minimum-Volume Enclosing Ellipsoids. *Optimization Methods & Software,* Vol. 23, No. 1, 2008, pp. 5–19.
36. Canon, M. D., C. D. Cullum, and P. Wolfe. On Improving Rate of Convergence of Frank–Wolfe Algorithm. *Siam Review,* Vol. 10, No. 4, 1968, pp. 470–498.
37. Daganzo, C. F. Queue Spillovers in Transportation Networks with a Route Choice. *Transportation Science,* Vol. 32, No. 1, 1998, pp. 3–11.
38. Schultz, G. W., and W. G. Allen, Jr. Improved Modeling of Non-Home-Based Trips. In *Transportation Research Record 1556,* TRB, National Research Council, Washington, D.C., 1996, pp. 22–26.

The opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

The Transportation Demand Forecasting Committee peer-reviewed this paper.

Will It Work?

Pilot Results from First Large-Scale Global Positioning System–Based Household Travel Survey in the United States

Greg Giaimo, Rebekah Anderson, Laurie Wargelin, and Peter Stopher

The Greater Cincinnati Household Travel Survey (HTS) is a proof-of-concept study for replacing travel diaries with a large-scale multi-day Global Positioning System (GPS) survey. The objectives are to collect multiple-day data from more than 3,000 households with portable GPS devices and improve existing processing software to provide data that support modeling approaches in Ohio. No diaries are collected for household members younger than 12 years old. A subsample of follow-up prompted recall surveys allow respondents to review GPS interpreted travel information for verification. This paper, with data from the spring 2009 pilot, describes the survey process developed for this HTS. It documents that with an address-based sample frame, advance letters, and Internet and phone recruiting, a significant subsample of cell phone–only households can be recruited and surveyed with GPS; a representative sample of households can be recruited for a GPS-based survey, based on a comparison of pilot sample household characteristics with available Public Use Microdata Samples data; and response rates for difficult-to-reach households such as cell phone–only, lower income, and zero-vehicle households can be improved with a cash incentive ($25). The paper provides principles and describes the prompted recall survey developed to obtain additional data from a subset of respondents beyond the GPS recorded travel for improving imputation software.

The first large survey based on the Global Positioning System (GPS) and conducted in the United States is the Greater Cincinnati Household Travel Survey for the Ohio city. Beyond various logistic issues, it is uncertain to what extent a 100% GPS-based survey can capture all the information available in a diary-based survey. This paper describes the processes used for the spring 2009 GPS survey pilot and presents the preliminary findings. The results of the recruitment and response by households, the resulting data, and the analyses are documented with a view to understanding whether a fully representative sample of households can be analyzed using solely GPS devices to collect travel information for all household members older than 12 years old—without the aid (and burden) of respondent-recorded trip diaries.

G. Giaimo and R. Anderson, Office of Multimodal Planning, Ohio Department of Transportation, 1980 West Broad Street, Columbus, OH 43223. L. Wargelin, Abt SRBI, Inc., Michigan Office, 19880 Maxwell Street, Northville, MI 48167. P. Stopher, PlanTrans Pty Ltd., 2 Wollemi Road, Woongarrah, New South Wales, Australia. Corresponding author: L. Wargelin, l.wargelin@srbi.com.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 26–34.
DOI: 10.3141/2176-03

BACKGROUND

Household travel surveys (HTSs) are designed to provide information about daily travel patterns, including trip purposes, time-of-day decisions, mode choices, trip lengths and distances, activity locations, and routes taken. This information is typically gathered from self-reported information in a diary, which is most often used as the basis for retrieving information with a computer-assisted telephone interview (CATI). Unfortunately, it is well documented that self-reporting leads to inaccuracies in travel information. The biggest shortcoming is (arguably) trip underreporting. For instance, recent work by Wolf et al. (*1*), Bradley et al. (*2*), and others comparing a subset of diary-reported household travel with GPS-recorded data has shown that diary information retrieved through CATI suggests trip underreporting ranging from 20% to 30%, with a worst case of as much as 60%.

In addition to the failure to report the number of trips correctly, comparative CATI versus GPS data collection procedures have found that respondents tend to overestimate trip times and underestimate (seriously) travel distance. Nonmotorized travel (particularly walking) is also thought to be poorly recalled compared with motorized travel, although the extent of this discrepancy has not been established scientifically. Location information tends to be even more problematic, with people rarely able to provide address information for even commonly visited destinations such as work, school, and the local grocery store to the degree of specificity required for geocoding and planning purposes (*3*). The situation is even more accentuated when trying to determine the route taken, with few people able to detail the route taken in terms of a sequence of street names.

An additional perceived problem with diary-based approaches and any type of phone or Internet survey of this nature is the burden on the respondent. The burden increases as the detail required and the number of days of observation increase. While most HTSs are 1- or 2-day surveys, evidence suggests that extending the survey for 3 days or longer results in greater statistical efficiency (*4*).

These issues aside, arguably the most pressing problem faced by all surveys is nonresponse. While there is marked variability depending on the strategies used, as a rule one can anticipate a 20% to 30% response from a mail-back survey, 25% to 40% percent from a telephone survey, and 60% to 75% percent for a face-to-face interview. However, nonresponse rates are not evenly distributed across the population, with certain groups (teenagers, larger households, and people who travel more) underrepresented in surveys (*5*). This situation leads to the potential for significant bias, which can be accounted for only partially in post-survey weighting of data results. With the many recent developments in improving the capabilities

and user-friendliness of small portable GPS devices, the time appears ripe to test the potential for GPS to replace travel diaries.

SAMPLING PLAN

The Greater Cincinnati Household Travel Survey is being enhanced further by recruiting from address-based samples within the eight-county survey study area rather than from traditional random-digit-dial (RDD) sampling frames. When all sources of undercoverage in RDD frames (i.e., households with no telephones, those in zero blocks, and cell phone–only households) are considered, the percentage of U.S. households not covered by RDD frames may be as high as 30%, depending on the metropolitan area. The survey households are being sampled from the U.S. Postal Service Delivery Sequencing File (DSF) frame, which is based on residential housing unit addresses (commonly referred to as deliverable residential addresses). The frame includes city-style addresses and post office boxes and covers single-unit, multiunit, and other types of housing structures. Known business addresses are excluded. A national survey sample vendor provides access to the DSF file and conducts the sampling to the project's specifications. By sampling U.S. Postal Service addresses instead of using random telephone numbers, the survey can focus on more specific areas. This practice improves the geographic representativeness of the sample and allows for oversampling of hard-to-reach but interesting household groups, such as those in transit-oriented neighborhoods, while maintaining the ability to expand the sample to be representative of the population as a whole by using differential expansion weights.

The sampling plan for the main survey provides for 3,000 to 3,600 households completing GPS-based travel inventories, with 700 to 1,000 households completing prompted recall (PR) surveys for additional information over a continuous time frame of 1 year. A complete household requires that every person have a complete GPS record or diary on a single travel day for one- to three-person households, and every person but one for four-person (or more) households. A complete PR survey consists of one person from a household completing the PR survey. Every recruited household member more than 12 years old carries a personal GPS unit for 3 consecutive assigned travel days. No diary recordings are substituted for members more than 12 years old. A simplified activity diary for children under age 13 is provided. For the main data collection effort, 500 households per month are recruited. Census block groups within the region with a higher transit propensity and near universities are being oversampled. Regional household characteristic distributions are monitored independently for specific targets by number of automobiles by number of workers, number of automobiles by household size, income categories, and life cycle (household type) using Public Use Microdata Samples (PUMS) data. Specific household characteristic distributions are expected to be met each month. The three geographic sampling areas are controlled on a quarterly basis. Distributions by political jurisdiction (counties and states) are also monitored on a quarterly basis.

This paper reports the results of the pilot test with 100 completed households (and 30+ GPS–PR surveys) conducted during spring 2009. For the pilot, census block groups with higher transit propensities were oversampled. It was expected that this oversample would provide what is needed in terms of overall household characteristic distributions (especially for households with adult students, low-income residents, and zero automobiles).

SURVEY PROCESS

Once the sample is selected according to specifications, the sampling vendor attaches land-line phone numbers to the addresses in the selected sample. A phone number can be matched with an address for about 55% of the sample. These addresses are called "matched samples." The "unmatched samples" without phone numbers are households with unlisted land-line phone numbers, no phones, or cell phones only.

The developed survey process has a two-pronged approach for recruiting households for the GPS-based HTS, depending on whether the sample household is within the matched sample or the unmatched sample stratum. Figure 1 shows this process flow.

Because addresses are available for all sample households, all households are sent advance letters explaining the survey and each household receives a password and directions for accessing the project recruitment interview on the web. For the pilot, those households without phone numbers (the unmatched sample) were also given a short "hot button" transportation issue questionnaire with questions such as: "Do you think it is more important to build a light rail system or focus on highway improvements in the Greater Cincinnati region?" Households were asked to complete the survey and return it (in a postage-paid return envelope) with their phone number for participation in the GPS-based survey (if they did not complete the recruitment online). Respondents had the option of calling a 1-800 number to provide a phone number.

After 1 week, if a sample household has not completed the recruitment interview online and a phone number is available either from the matching process or from postcard returns or 1-800 call-ins, the household is called (as in an RDD survey) to complete the recruitment interview over the phone. During the 6-min phone or Internet recruitment interview, three consecutive travel days are assigned. Once recruited, GPS units and instructions, household and person forms, and children's diaries for those less than 12 years old are deployed. The forms, shown in Figure 2, collect work and school locations, the two most frequent household shopping locations, and GPS usage status for each member on each day. A reminder phone call or e-mail takes place the day before the travel date.

The GPS devices used are personal units (smaller than a small cell phone) that can be carried in a pocket or purse or clipped on a belt or wrist band. Thus, they record all modes of travel including car, transit, bike, and walk. For the most part, the units record 3 days of travel. In the pilot survey, respondents were not provided with battery chargers and some respondents apparently toyed with the devices so that batteries ran out much sooner. In the main survey, battery chargers are supplied and respondents are encouraged to charge the units each night.

Finally, a goal of this large-scale GPS-based HTS is to develop an efficient (low-cost) means of deployment of the units to and from widely scattered sample households around the metropolitan region, because the costs of full personal courier delivery and collection are prohibitive. The survey procedure plan is to send out the GPS units and forms by Federal Express (at a government rate of about $8 per package). The outgoing package contains prepaid return shipping labels and a return package that can be deposited in any Federal Express or U.S. Postal Service drop box. Household respondents are also given the project 1-800 number to arrange a Federal Express or personal courier pickup if they prefer. Extensive follow-up phone calls and Internet reminders help arrange courier pickups as needed, as the second most expensive and difficult logistics challenge presented by GPS-based surveys can be excessive loss of GPS units.

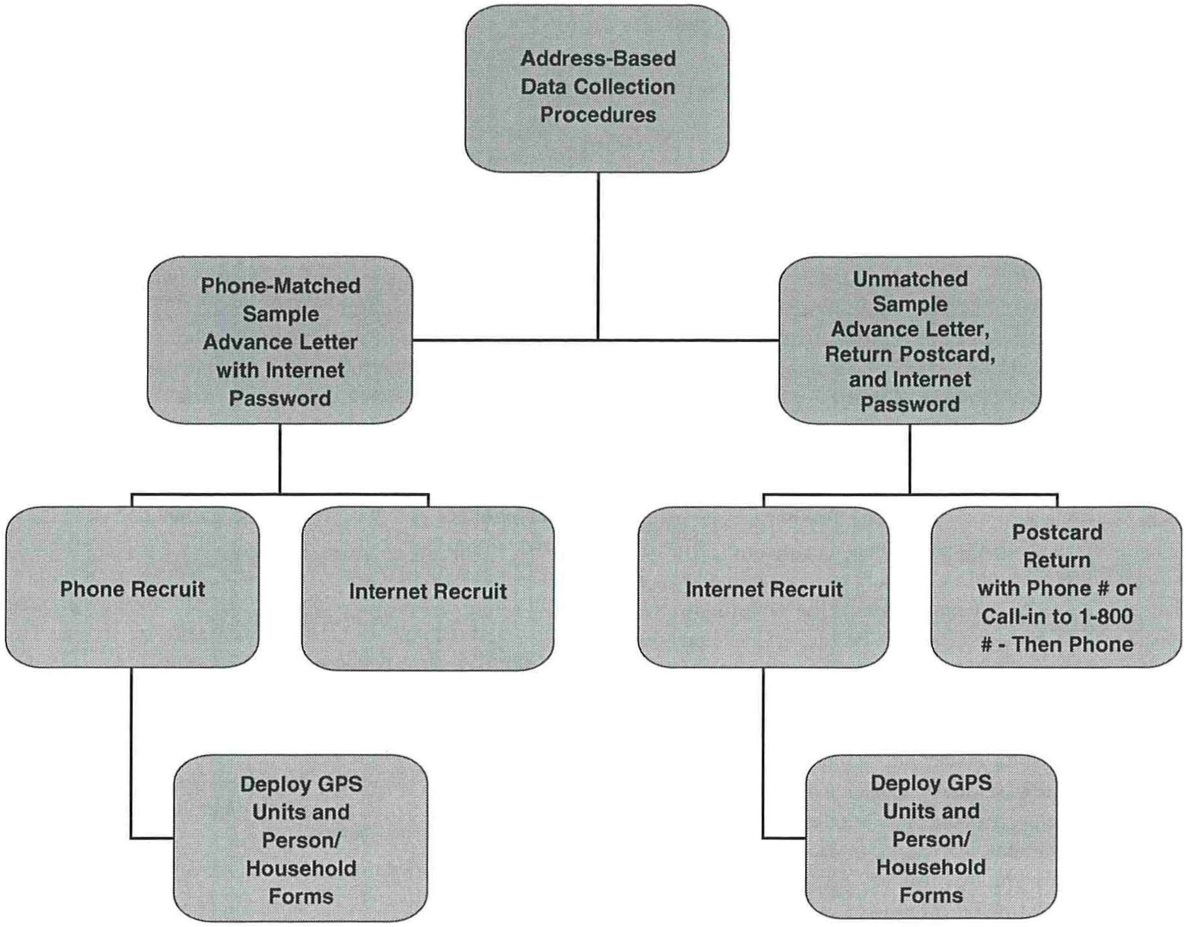


FIGURE 1 Address-based sample recruitment procedures for GPS-based HTS.

PILOT OBJECTIVES

The research team sought to obtain an early understanding of cooperation and completion rates for different survey respondent groups and to test the potential effectiveness of respondent incentives. As shown in Table 1, the pilot was designed to explore differences in response rates among five groups:

- Respondent households within and outside the transit propensity area,
- Respondent households in the matched address–telephone number sample and unmatched address–telephone number sample,
- Completion incentives of $0 and $10 for the matched address–telephone number sample,
- Completion incentives of $10 and $25 for the unmatched address–telephone number sample, and
- Completion rates for all key household characteristic segments.

For the pilot, researchers sought the ability to analyze response rates for different realistic combinations of treatments. Thus, the sample deployed had the following characteristics:

- Equal sample for higher transit propensity and medium-to-lower transit propensity households,
- Equal sample for phone-matched and unmatched sample,
- Equal proportion of matched sample offered $0 or $10 incentive to complete, and
- Equal proportion of unmatched sample offered $10 or $25 incentive to complete.

Allocating the pilot sample in this way enabled a statistical comparison of the outcomes between the transit and nontransit areas, the matched and unmatched samples, and different incentive levels. All respondent households that completed the GPS portion were recruited for the PR survey.

PILOT RESULTS

Overall, the ratio of completes to recruitments was relatively stable for all geographic and demographic segments. In other words, the makeup of the household sample that completed the GPS survey was similar to the sample that was recruited in terms of both matched and unmatched sample and higher and lower transit propensity areas.

As shown in Figure 3, there was no difference in response and completion rates for those offered $0 and $10 incentives among the matched sample (with phone numbers). Among the unmatched sample (mostly cell phone–only survey households) a $25 incentive doubled completion rates.

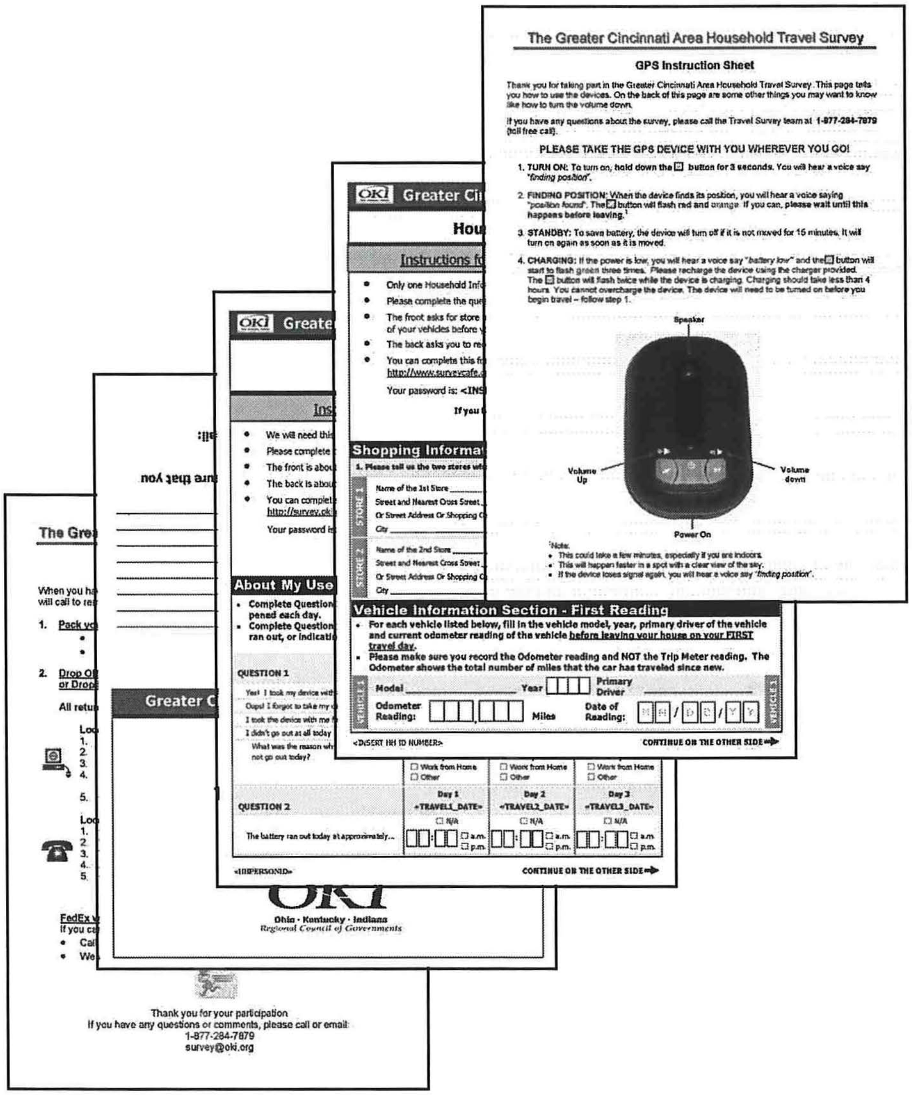


FIGURE 2 Short-form materials piloted with deployment GPS units.

As Figure 4 shows, in line with the pilot sampling plan, nearly equal percentages of pilot GPS completes were from transit and nontransit areas—there was little difference in participation rates. However, of total completes, 81% were from the matched phone sample, while 19% were from the unmatched sample (cell phone or unlisted number only). As shown in Figure 5, in relation to the incentives, a $10 incentive versus no incentive made little difference among households with less than $50,000 incomes and only a slight difference for those with $50,000+ incomes. Offering a $25 incentive made a significant improvement in the percent of completes among all income groups. The $25 increased incentives among households with incomes greater than $50,000 as well as among lower-income households, but without incentives the representation of these households in the completed sample was sufficient. Therefore, it can be concluded that significant incentives are needed to improve completion rates among households with less than $50,000 incomes.

TABLE 1 Pilot Sample Design

Matched Sample	Unmatched Sample
Higher transit access subsample $0 incentive	Higher transit access subsample $10 incentive
Lower transit access subsample $0 incentive	Lower transit access subsample $10 incentive
Higher transit access subsample $10 incentive	Higher transit access subsample $25 incentive
Lower transit access subsample $10 incentive	Lower transit access subsample $25 incentive

With respect to the matched and unmatched sample, the Internet was the most effective means of obtaining recruits from households without land-line phones or listed phone numbers. Additionally, 19% of recruits from the phone-matched sample responded to the advance letter by completing the recruitment online. Only one phone number was obtained from the unmatched sample via a return postcard or reply to a hot button issue survey. Therefore, this option was eliminated for the main survey. Regardless of recruitment method, once recruited, completion rates for the matched and unmatched sample were equivalent.

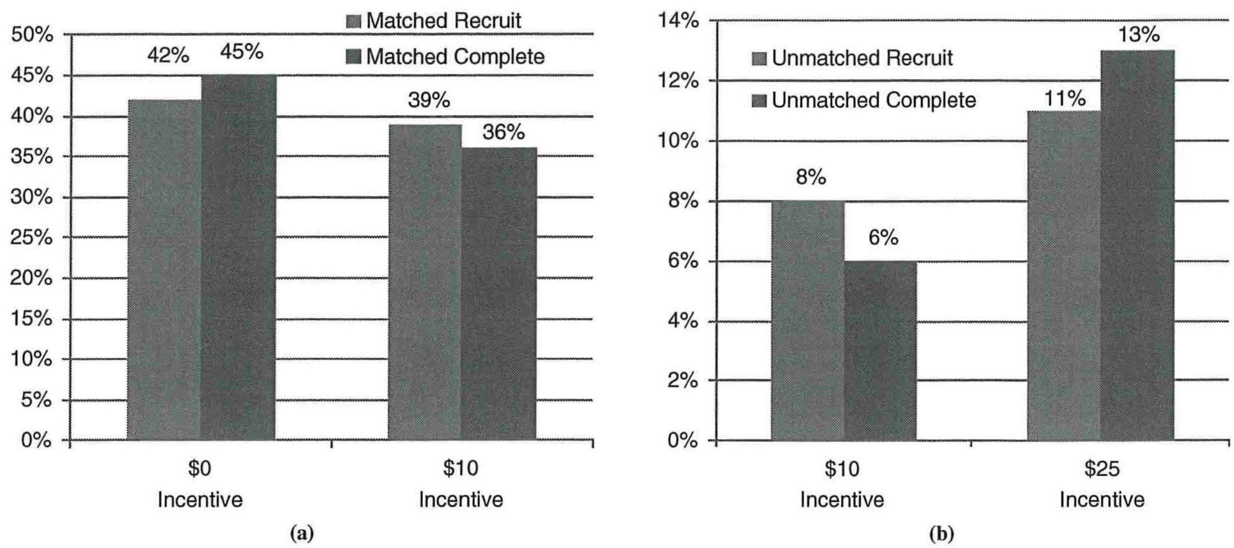


FIGURE 3 Completion to recruitment rates for sample segments.

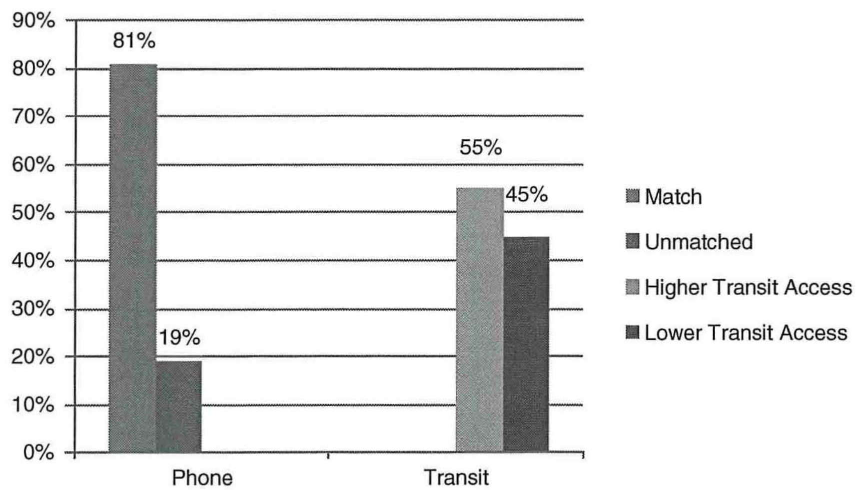


FIGURE 4 GPS completions as percent of total for matched and unmatched sample and for higher transit and lower transit areas.

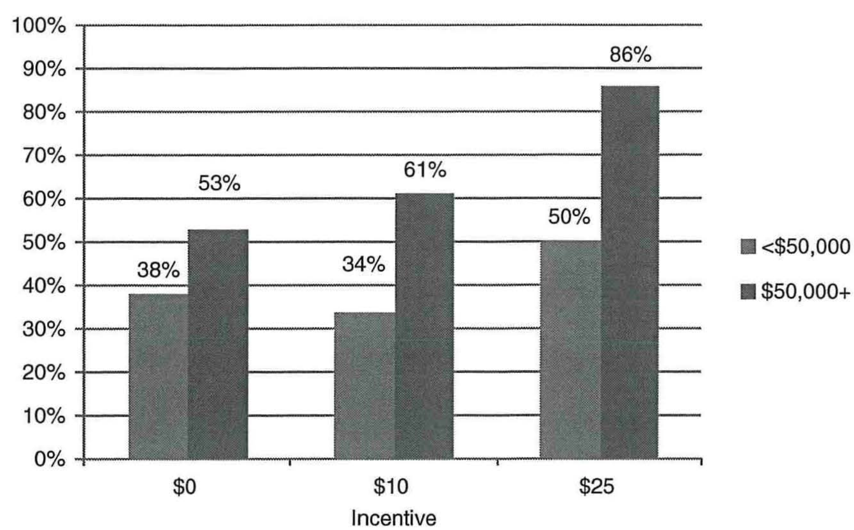


FIGURE 5 Differential impacts of incentives on income groups.

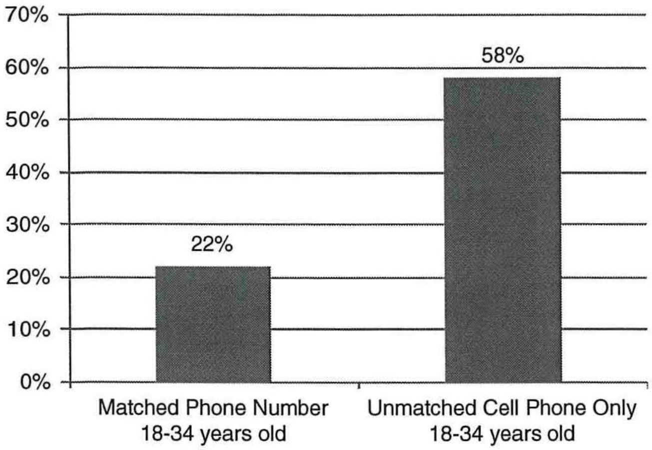


FIGURE 6 Contact person (18 to 34 years old) for matched phone sample and unmatched cell phone–only sample.

Respondents 18 to 34 years old are typically underrepresented in diary HTSs, and comparisons of diary-recorded trip inventories with subsamples of GPS surveys show their trips and tours to be underreported. Figure 6 shows that this underreporting age group (who are most likely to be cell phone–only households) was captured for the GPS-based HTS through an advance letter and go-to web-based recruitment interview.

Tables 2 through 6 show the breakdowns of completed households by the key household characteristics compared with the latest available PUMS data. With respect to these key household characteristics, the pilot survey showed that a representative sample was recruited and completed by household size. The requirement that all household members age 13 or older carry GPS units did not prove to be a "respondent burden" barrier. A representative sample was also completed by household number of vehicles. However, while an appropriate percentage of zero-vehicle households were recruited, extra effort in the form of incentives ($25) will be needed to complete zero-vehicle and low-income households. Tables 5 and 6 show that nonworking households were easier to recruit than households with workers but that it was difficult to recruit low-income households to participate. Again, these differences are commonly seen in all HTSs and are not unusually present in this GPS survey.

PILOT SURVEY PROCESS LOGISTIC ISSUES

From the pilot survey, a number of logistic issues that have both timing and cost implications for the main survey were found. Retrieving GPS units in a timely manner for redeployment was an underestimated problem. The loss rate for the pilot was 2.7%, mostly among low-income, urban households. More units are needed for the main survey than expected to cover this ongoing estimated loss rate. Incentives of $25 are being offered to the unmatched sample, to households with incomes below $25,000, and to zero-vehicle households upon completion of the GPS survey and return of the GPS device. Some battery outages were experienced over the 3 assigned travel days, resulting in some households without a complete travel day recorded for each member more than 12 years old. In the main survey, battery chargers and instructions are being supplied to households.

KEY ELEMENTS OF SURVEY DESIGN AND PROMPTED RECALL INTERVIEW

Because the GPS survey was designed to replace diary-based surveys, additional information was needed that is typically revealed in a diary-based survey. For the GPS survey, in addition to the standard demographic and work and school location questions, two main shopping destinations are also requested for each household. An additional question is asked of each person assigned a device as to

TABLE 2 Representativeness of Pilot Sample by Household Size

Household Size	Percentage of Households	Percentage of Household Completes	PUMS 2000 Data (%)	Difference (%)
1	29.2	29.3	27.3	2.0
2	31.1	34.1	32.0	2.1
3	15.3	15.9	16.6	–0.7
4 or more	24.4	20.7	24.1	–3.4
Total	100.0	100.0	100.0	

TABLE 3 Representativeness of Pilot Sample by Automobiles Available to Household

Autos Available to Household	Percentage of Households	Percentage of Household Completes	PUMS 2000 Data (%)	Difference (%)
0	9.1	2.4	9.7	–7.3
1	28.2	28.0	32.3	–4.3
2	40.2	46.3	38.8	7.5
3 or more	22.5	23.2	19.2	4.0
Total	100.0	99.9	100.0	

TABLE 4 Representativeness of Pilot Sample by Life Cycle

Household Type	Percentage of Households	Percentage of Household Completes	PUMS 2000 Data (%)	Difference (%)
Adult household	48.3	43.9	46.0	–2.1
Household with children	32.5	34.1	36.6	–2.5
Retiree household	14.8	18.3	14.5	3.8
Adult student household	4.3	3.7	2.9	0.8
Total	99.9	100.0	100.0	

TABLE 5 Representativeness of Pilot Sample by Number of Workers in Household

Number of Workers in Household	Percentage of Households	Percentage of Households Completes	PUMS 2000 Data (%)	Difference (%)
0	29.7	32.9	24.0	8.9
1	37.8	35.4	37.4	–2.0
2	26.8	30.5	31.3	–0.8
3 or more	5.7	1.2	7.3	–6.1
Total	100.0	100.0	100.0	

TABLE 6 Representativeness of Pilot Sample by Income

Income Category	Percentage of Households	Percentage of Households Completes	PUMS 2000 Data (%)	Difference (%)
Less than $25,000	21.8	9.1	20.6	–11.5
$25,000 to less than $50,000	26.9	28.6	25.1	3.5
$50,000 to less than $75,000	19.3	20.8	20.2	0.6
$75,000 or more	32.0	41.6	34.1	7.5
Total	100.0	100.1	100.0	

its use for the day, whether the device was carried all day, whether it was forgotten for part of the day, whether the battery died during the day, or whether the device was forgotten the entire day.

To provide data to improve the software for mode and purpose identification, a PR survey was also implemented. It shows each respondent's travel on the computer screen and asks a series of questions about that travel, such as mode and trip purpose. The PR responses are then used to improve software, which will be used to impute trip mode, purpose, and other missing data for the other completed surveys.

GPS Data Imputation and Verification: Processing Methodology

The software used to process the GPS data was developed at the University of Sydney, Australia, and is described in detail elsewhere (*3*). The first of these programs uses a number of rules to delete spurious data (generally the data collected while the device is at rest at the end of a trip or possibly in the middle of a trip when there is a lengthy delay in movement, such as at a traffic signal) and to split the data stream into what are assumed to be individual trips. At the completion of this process, maps are generated by the software, along with a summary file showing the assumed start and end locations of each trip, the time (to the nearest second) when the trip started and ended, and some other characteristics of the trip (e.g., distance, elapsed time, average speed).

Because it is not possible to craft rules that will work 100% of the time, the next step in the process is "map editing." Trained staff review the maps for each day using geographic information system software and look for possible spurious data that may not have been deleted in the initial processing, for possible stops in a trip that the software identified as a single trip, and for trips that might be split into two or that may be missing because of loss of the GPS signal. Trips may not be split correctly by the software because of a rule that dictates that an identifiable stop (after spurious data are removed) must last at least 120 s to define the end of a trip. Because a number of activities take less than 120 s to accomplish that should also define the end of a trip (such as picking up or dropping off a person), and also because the deletion of spurious data is done conservatively by the software, it is necessary to inspect the map and make some edits to the list of trips provided by the software processing.

After map editing is complete, the data are run through several processes before producing the data for the web survey. One process applies the changes from the amended trip list file to the original trip database to remove data points and to split or join trips. Another process compares the address information collected from respondents with the locations of trip ends in the modified trip list and records matches for input into the web survey, so that home, work,

educational establishment, and grocery shopping locations can be shown on the map and the possible purpose of the trip can be shown. An Internet address is generated for each individual respondent, with the edited GPS trips and asking questions of the respondent about these trips, for the purposes of verifying the software processing and providing inputs for further software refinements. The data are also processed to identify mode and purpose, using a number of heuristics that are described elsewhere (*5*). The information deduced by this software is not included in the web survey but is compared with what respondents indicate in the web survey.

The PR Internet format went through several revisions during the pilot phase. About 27% of households completed the PR, but there was a fall-off among larger households. The purpose of the PR is for the respondent to provide feedback on the GPS interpreted travel information for the assigned and recorded days. Figure 7 provides a screen view of the PR in its pilot form.

Basic Design Principles for Prompted Recall

The GPS devices record all location data. The only errors (provided that all downloading and processing are done correctly) that can occur in the location data are as follows:

1. Cold start problem. The device is delayed in fixing position until after a trip has started. This problem is fixed by the software for all except the first trip on the first day and the first trip on the first day is fixed by the map editing that precedes setting up the Internet address for each respondent to the PR survey.

2. Lost signal. Signal loss is a problem only if it occurs near the end of a trip and results in a premature destination recording. This problem should normally be fixed in the map editing process before setting up the PR survey.

The other things that can be wrong with the GPS record are that people did not carry the device with them all day and that the battery ran down. In these cases, if people marked on their form that they forgot the device for some of the day or that the battery ran down, then that day is excluded from the sampling for PR. Apart from these issues, the start and end times of travel on the GPS record must generally be correct, and, provided that the travel is also along the street or rail networks, then a trip must have taken place.

Possible Errors and Omissions in GPS Record

This process leads to the assumption that the only things that can be wrong with the GPS record for PR respondents that necessitate editing the data are as follows:

1. GPS processing has missed identifying a brief stop (usually one that lasted less than 120 s). In this case, the respondent should be allowed to insert one or more stops, thereby splitting one trip into two or more trips.

2. GPS processing has identified as a stop that was actually just a traffic stop or other delay that lasted at least 120 s. In this case, the respondent should be allowed to delete one or more stops, thereby linking together two or more travel episodes.

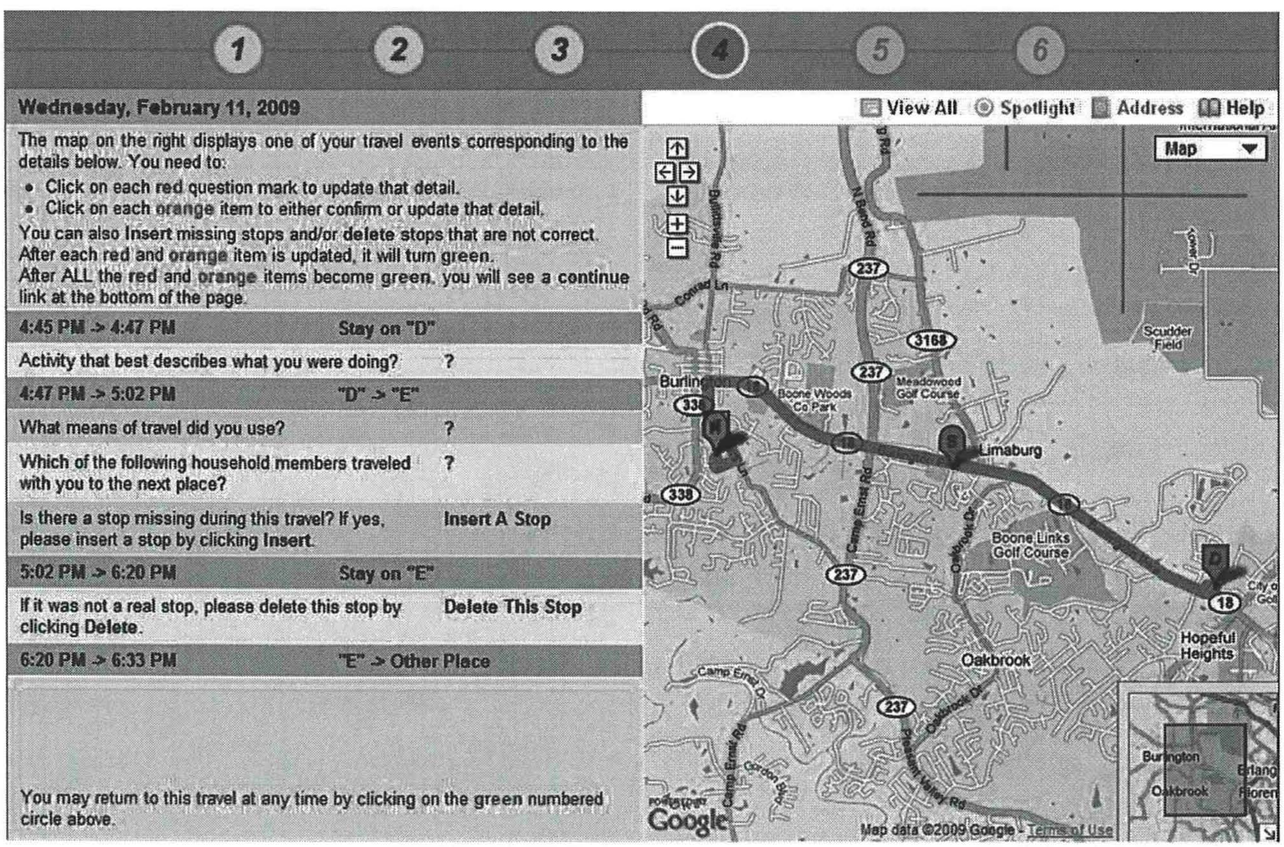


FIGURE 7 PR web format.

Proposed Design for Deletion or Addition of Stops

With these assumptions, questions of confirmation of the start and end times of travel are not asked in the main PR survey, although they were in the pilot survey. First, to allow for deletion of a stop, a stop is displayed to the respondent, followed by travel, followed by the next stop, followed by the next travel. Because the respondent can see and edit the second stop and can see the second travel event, then the respondent can have the option of deleting the second stop, thereby joining the first and second travel event.

Second, a question is included to ask whether the person traveled from the first to the second stop without stopping. If the person responds no, then an edit box pops up that allows the individual to insert the time stopped and the time travel started again. This process also automatically updates when confirmed by the respondent and will shorten the display to show two stops and two travel events. After completion of all the other travel details (mode, companions) and the nature of the stops, the respondent will click on continue, which will display the second stop, the second travel event, the third stop, and the third travel event of the day. This process continues in pairwise fashion to the end of the day.

If the respondent indicates reasonably accurately the time stopped for each added stop, then the GPS record (with speed) will allow identifying where the additional stop was with reasonable accuracy.

A closing PR question will ask whether there is any travel or any other stop that the respondent remembers making on that day that could not be recorded on the survey. If yes, the individual will be asked to record stops and approximate times. The PR survey is then complete.

CONCLUSIONS

The pilot of the Greater Cincinnati GPS-Based HTS demonstrated the following:

- Address-based sampling can be successful in recruiting cell phone–only survey households to a GPS-based HTS with an Internet recruit. Once these households are recruited, they complete the GPS HTS in similar proportion to other households.
- A fully representative sample by characteristics of household size, number of automobiles, number of workers, income, life cycle, and geographic region can be completed by a GPS-only HTS, recording up to 3 days of travel.
- GPS household completion rates are adequate as well as representative.
- Requiring every household member (older than 12 years) to carry a GPS unit for 3 days was not considered an undue burden; paperwork and telephone recruit and retrieval times were greatly reduced.
- Significant incentives and additional efforts are needed to complete address–phone number unmatched households and households with low incomes or zero vehicles.
- Added trip accuracy reporting and value of route and location with speed data (collected with GPS) need to be demonstrated upon completion of pilot and main survey trip files.

REFERENCES

1. Wolf, J., M. Loechl, M. Thompson, and C. Arce. Trip Rate Analysis in GPS-Enhanced Personal Travel Surveys. In *Transport Survey Quality and Innovation* (P. Jones and P. Stopher, eds.), Elsevier, London, 2003, pp. 483–498.
2. Bradley, M., J. Wolf, and S. Bricka. Using GPS Data to Investigate and Adjust for Household Diary Data Non-Response. Presented at 10th Transportation Planning Applications Conference of the Transportation Research Board, Portland, Ore., 2005.
3. Stopher, P., E. Clifford, C. FitzGerald, and J. Zhang. Deducing Mode and Purpose from GPS Data—Case Studies. Presented at 11th Transportation Planning Applications Conference of the Transportation Research Board, Daytona Beach, Fla., 2007.
4. Wolf, J. Application of New Technologies in Travel Surveys. In *Travel Survey Methods—Quality and Future Directions* (P. Stopher and C. Stecher, eds.), Elsevier, Oxford, U.K., 2006, pp. 531–544.
5. Stopher, P., M. Xu, and C. FitzGerald. Assessing the Accuracy of the Sydney Household Travel Survey with GPS. *Transportation,* Vol. 34, No. 6, 2007, pp. 723–741.

The Transportation Planning Applications Committee peer-reviewed this paper.

Internet Access

Is Everyone Online Yet and Can We Survey Them There?

Colin Smith and Greg Spitz

Travel surveys can reach potential survey respondents in many ways, including intercepting them at activity centers or on transit vehicles and contacting them by telephone, mail, or online. The selection of a survey recruitment approach should consider whether the population from which a sample is drawn is fully covered in order to avoid coverage error. This paper presents two case studies that review data collected in travel surveys in metropolitan areas in the United States to identify coverage error, specifically whether lack of Internet access by a segment of the population leads to coverage error. The first case study analyzes data from several road pricing surveys to quantify differences between those who have and do not have Internet access. The second case study analyzes data from two transit origin–destination surveys in which respondents were invited to provide contact details for additional market research. This paper compares the overall sample with those willing to be surveyed in the future by telephone and those willing to receive a future survey invitation by e-mail. Both case studies find that the samples with access to the Internet are similar to the larger full samples that include those without Internet access and therefore that the coverage error found in the Internet-only samples is small. The results suggest that, for surveys of general populations of drivers or transit riders, surveying only those with Internet access does not introduce significant coverage error into travel survey samples.

Travel surveys can reach potential survey respondents in many ways, including intercepting them at activity centers such as shopping malls, employment centers, and public buildings; at the roadside, at toll plazas, or on transit vehicles during the course of travel; by telephone or mail to residents of a defined study area; and online through e-mails to an Internet survey panel or customer list.

The availability of computer-based survey instruments that allow surveys to be taken online has led to the following research question: is coverage error introduced if only respondents who have Internet access are recruited to complete a travel survey? Dillman describes coverage error as, "the result of not allowing all members of the survey population to have an equal or known chance of being sampled for participation in the survey" (*1*, p. 11). In the context of this paper, the populations that may be sampled for travel surveys include transit riders, toll road customers, potential users of a planned road or transit system, and a sample of all travelers in a given geography. Coverage error is of concern as the omitted segment that is not covered may differ in some important way from the covered segments of the population, possibly leading to different inferences being made based on the survey data than would occur had the sample been representative. This paper presents two case studies that review data collected in recent travel surveys to look for evidence of coverage error, specifically whether lack of Internet access by a segment of the population leads to coverage error in Internet-only samples. The second case study also compares the amount of coverage error in an Internet-only sample with the amount of coverage error in a telephone sample.

Resource Systems Group, Inc., 55 Railroad Row, White River Junction, VT 05001. Corresponding author: C. Smith, csmith@rsginc.com.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 35–41.
DOI: 10.3141/2176-04

SURVEYING BY TELEPHONE

An example of a survey sampling approach with growing coverage problems is random-digit-dial (RDD) telephone surveys. Recent research has highlighted the rapidly increasing proportion of households without landline telephones, as many households stop landline service and use only cell phones. Tucker et al. (*2*) explained that changes in the U.S. telephone system, especially the rapid growth in the prevalence of cell phones, raise concerns about coverage error in RDD telephone samples.

Blumberg et al. (*3*) found that cell phone–only households made up 14.7% of U.S. households in 2007 and ranged from 5% in Vermont to 26% in Oklahoma. Blumberg and Luke (*4*) also found that coverage error is particularly apparent for low-income adults, young adults, and most clearly for low-income young adults, in which the proportion of cell phone–only households was 32% in 2006.

Attempts are being made to supplement RDD samples with cell phone samples. Link et al. (*5*) found that conducting surveys by sampling cell phone numbers is feasible but is costly and produces relatively low rates of participation. They also confirmed that significant differences exist between cell phone–only households and those with other types of telephone access.

The literature on telephone surveys suggests that coverage error exists in telephone surveys that use RDD samples and that removing it is difficult and expensive. Despite the well-documented presence of coverage error, telephone surveys continue to be used for many travel surveys, and in many cases they are conducted without using an additional cell phone sample.

SURVEYING ONLINE

Couper and Miller explained that

> while we have seen dramatic growth in internet access and use in both the developed and developing world, the proportion of those with internet access appears to have reached a plateau in the USA in recent years. Further, regardless of the actual level of penetration, substantial proportions of the population remain without access to the internet, and

the differences between those with and without access show no signs of diminishing. (*6*, p. 832)

The Pew Internet and American Life Project published research (*7*) that shows that the percentage of U.S. adults online stabilized between 70% and 75% from the middle of 2005 until the end of 2008, but during 2009 it increased by a few percentage points to 77%. Despite the recent increase, Pew's findings suggest that a significant segment of the population will not have Internet access in the immediate future.

Other research by the Pew Internet and American Life Project shows that Internet use is more prevalent among younger age groups, residents of urban areas, higher-income groups, and groups with higher levels of educational attainment (*8*). Among the youngest age group of adults, 18- to 29-year-olds, 93% have Internet access, compared with 83% of 30- to 49-year-olds, 77% of 50- to 64-year-olds, and 43% of those 65 years old or older. Rates of Internet access drop from 95% of those with household income above $75,000 to 62% of those with household income under $30,000. For those with a college education, the rate of Internet access is 94%; this share falls to 37% for those with less than a high school education.

These data suggest that conducting a travel survey that sampled only those with Internet access could lead to coverage error. Certain demographic groups with lower levels of Internet access, such as those with lower incomes and older populations, may exhibit different travel behavior than groups with higher levels of Internet access.

CASE STUDY 1. TOLL ROAD STATED PREFERENCE SURVEYS

The data used for the analysis presented in this case study are from surveys of highway users. The main objective of these surveys was to evaluate drivers' willingness to pay tolls or congestion charges for new roads, managed lanes, toll bridges, or central area congestion charges. This case study reviews the data to look for evidence of coverage error and for differences between those with and without Internet access.

Survey Approach

The data used for the analysis presented in this case study were collected in six areas around the United States during 2007 and 2008. Four of the six locations are in the southeastern part of the country, with two surveys in different regions in Florida, one survey in Atlanta, Georgia, and one survey in Jackson, Mississippi. The other two surveys are in San Francisco, California, and Anchorage, Alaska.

To gain representation from different groups of drivers, a multimethod sampling approach was used for each survey, where several recruitment methods allowed a diverse sample to be collected. For analysis, the various respondent recruitment methods were grouped into four categories:

1. Activity locations. Respondents are intercepted at activity locations close to the study corridor or area and asked to complete a computer-based survey on a laptop at a kiosk set up at that location. The locations include shopping malls, Department of Motor Vehicle offices, libraries, and public buildings such as city halls.
2. Employer e-mails. Large employers such as corporations, hospitals, and universities are contacted and asked to send an e-mail invitation to their employees.
3. Handout or mail-out. A postcard with an invitation to take the survey online is mailed to a sample of respondents living in a study area or handed out to a sample of respondents using a toll plaza or highway, or respondents to a handout or mail-out paper origin–destination survey are asked for their e-mail addresses on the survey and e-mailed an invitation.
4. Purchased sample. Members of online survey panels are invited to take the survey.

Table 1 shows the sample sizes for the six surveys grouped by data collection approach. The analysis presented in this case study is based on the 4,071 survey records collected at activity locations (second column in Table 1).

Each survey included a pair of questions asking whether the respondent had Internet access and, if so, whether it was at home, work, an "Internet café, library, or other public place using my own computer," or at an "Internet café, library, or other public place using their computer terminal." The surveys also asked respondents a set of demographic questions such as age, gender, household size, household vehicle ownership, and household income along with questions about their travel in the study area for the project being studied.

Survey Results

Internet Access

This case study reviews only data collected at activity locations. Data are collected at a set of activity locations in a study area that are chosen to ensure that the sample is as representative as possible;

TABLE 1 Sample Sizes for Each Survey Location by Data Collection Approach

Survey Location	Activity Locations	Employer E-mails	Hand–Mail Out	Purchased Sample	Total
Anchorage, Alaska	435	17	266	0	718
Atlanta, Ga.	1,812	1,278	117	966	4,173
Broward County, Fla.	310	45	247	200	802
Central Florida	519	469	0	429	1,417
Jackson, Miss.	331	221	733	0	1,285
San Francisco, Calif.	664	1,024	0	3,616	5,304
Total	4,071	3,054	1,363	5,211	13,699

TABLE 2 Type of Internet Access for Sample Intercepted at Activity Locations by Survey

Survey Location	Home (%)	Work (%)	Public Space with Own Computer (%)	Public Space Using Terminal Computer (%)
Anchorage	91	48	15	10
Atlanta	88	55	20	14
Broward County	93	40	17	14
Central Florida	81	33	16	17
Jackson	85	52	26	19
San Francisco	96	82	20	6
Total	89	55	19	13

for example, surveying might take place at several shopping malls including regular malls and "upscale malls." Generally, the sample collected at activity locations is from a wide cross section of the population, as evidenced from the responses to the demographic questions in the survey.

For this analysis, the data collected at activity locations required respondents merely to be intercepted at an activity location and meet the screening criteria for the survey. Respondents completed the survey at the activity location on a laptop computer provided by Resource Systems Group, Inc. The data collected at activity locations are the only category of data collected in these studies that did not require the respondent to have Internet access to take the survey.

The sample intercepted in downtown San Francisco has the highest rate of Internet access, with only 1% of respondents not having Internet access. In relatively affluent Broward County, Florida, only 7% of respondents did not have Internet access. In the remaining survey areas, the percentage of respondents who did not have Internet access increased to 10% in Atlanta and Jackson, 12% in Anchorage, and 16% in central Florida. Almost 90% of respondents who have Internet access have it at home (Table 2). More than half (55%) have Internet access at work and about one-fifth have Internet access in public spaces.

An alternative approach to reviewing the responses to this question is to build a hierarchy of Internet access. The first group in the hierarchy has Internet access at home and work, which amounts to half the respondents with Internet access. The second group, which amounts to 40% of those with Internet access, has access at home but not at work. For the remaining 10% of respondents with Internet access, 5% have access at work but not at home, 3% have no access at work or home but have a computer they can use for Internet access in a public space, and 2% are limited to using a terminal computer at a public space such as a library or Internet café (Table 3).

The sample from San Francisco stands out as having very high Internet accessibility, with 78% having access at both home and work. In no other survey location did the size of this group reach 50%. The results suggest that very small proportions of those who describe themselves as having Internet access rely on Internet access at work or in a public space. Only in central Florida did the proportion of respondents who rely on Internet access using a terminal computer exceed 2% of respondents with Internet access.

Demographics of Those With and Without Internet Access

A comparison of all respondents intercepted at an activity location who do not have Internet access with those who do have Internet access demonstrates some demographic differences between the groups. Those without Internet access are slightly more likely to be in single-person households: 24% of those with no Internet access live in single-person households compared with 17% of those with Internet access.

Those without Internet access are slightly more likely to be in the youngest age group (16 to 24 years) and the oldest age group (65 years or older) than those with Internet access (Figure 1). The age distribution of those with Internet access tracks very closely to the age distribution of the whole sample (this pattern is repeated for the demographics discussed below).

TABLE 3 Hierarchy of Internet Access for Sample Intercepted at Activity Locations by Survey

Survey Location	Home and Work (%)	Home, Not Work (%)	Work, Not Home (%)	Public Space with Own Computer, Not Home or Work (%)	Public Space Using a Terminal Computer Only (%)
Anchorage	42	49	6	2	1
Atlanta	49	39	6	4	2
Broward County	37	57	4	2	0
Central Florida	29	55	5	5	6
Jackson	48	41	7	4	1
San Francisco	78	18	4	0	0
Total	50	40	5	3	2

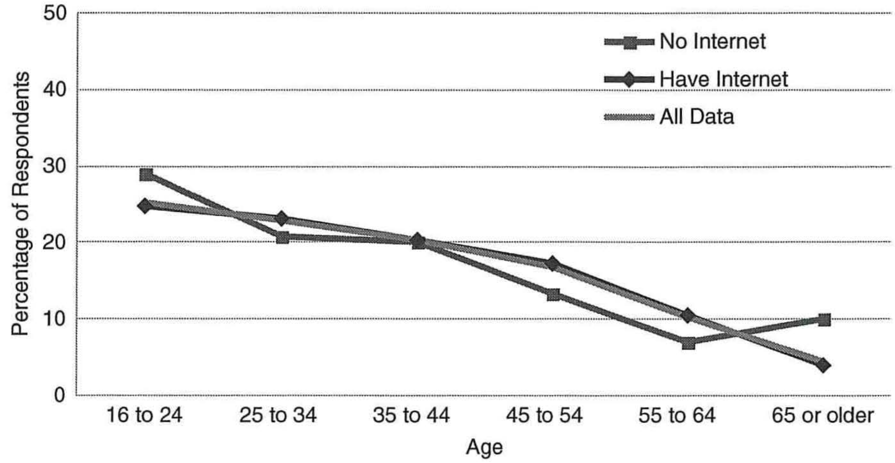


FIGURE 1 Age distribution by Internet access for sample intercepted at activity locations.

The screening criteria for the surveys analyzed in this case study require that respondents are driving in the study corridor or area and therefore practically require vehicle ownership. Therefore, the low rate of zero-vehicle households (3% or less) among respondents with and without Internet access is to be expected. However, those without Internet access are much more likely to be in one-vehicle households than are those who have Internet access (45% compared with 27% of respondents), and they are much less likely to be in households with three or more vehicles (22% compared with 34% of respondents).

Differences between those with and without Internet access exist in household income and employment status. Those without Internet access are more likely to be members of households with income below $50,000 than are those with Internet access (Figure 2). Those without Internet access are also more likely to be employed part time, retired, or not currently employed than are those with Internet access.

Case Study 1 Conclusions

In studies such as the six surveys presented here, in which respondents are screened for current or likely use of a road, the proportion of those with Internet access is slightly higher than in the general population. This effect is probably because the relationship between household income and Internet access also applies to vehicle ownership: low-income households are less likely to have Internet access and less likely to own a vehicle.

The results demonstrate that the sample with Internet access has a higher income and owns more vehicles than the sample without Internet access. The results also show that the sample with Internet access does not significantly differ in age, income, and vehicle ownership dimensions from the full sample.

The results suggest that, in general, collecting data only from those with Internet access is unlikely to lead to significant coverage error in a travel survey where driving in the study area is required.

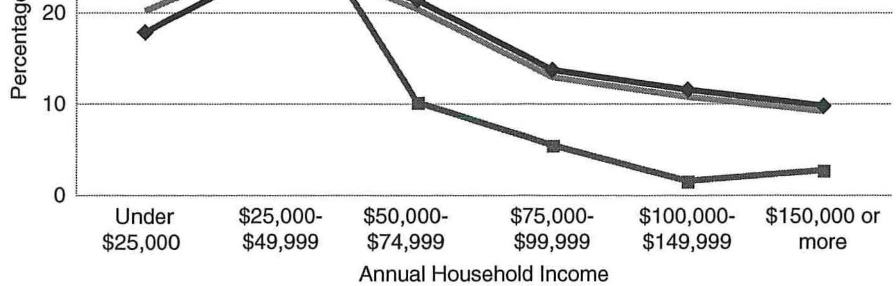


FIGURE 2 Household income by Internet access for sample intercepted at activity locations.

Omitting the small proportion of the driving population without Internet access would not change the demographic profile of the resulting sample. However, the results show that rates of Internet access vary by survey location. In areas where rates of Internet access are lower, there remains a possibility of coverage error. In those cases, approaches such as weighting the data to correct for undersampling of certain age and income categories or collecting data at activity locations to collect data from undersampled groups should be considered.

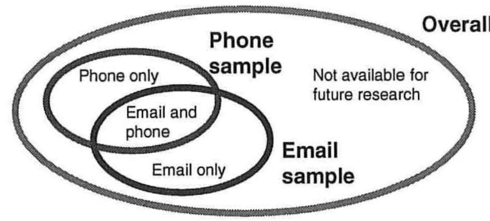


FIGURE 3 Transit origin–destination survey sample groups.

CASE STUDY 2. TRANSIT ORIGIN–DESTINATION SURVEYS

The data used for the analysis presented in this case study are from surveys of transit riders. The main objective of these surveys was to understand the origins and destinations of a representative sample of the population of transit riders. This case study reviews the data to look for evidence of coverage error and for differences between respondents willing to be contacted for future surveys by e-mail and those unwilling or unable to be contacted by e-mail.

Survey Approach

The data used for the analysis presented in this case study were collected in two locations during 2007: on the Chicago Transit Authority (CTA) bus and subway network in Illinois and on the Metro North Railroad (MNR) east of Hudson commuter rail network in New York and Connecticut. Recruitment was carried out on the transit vehicles in both surveys, with paper surveys handed out. Respondents could hand back completed surveys to survey staff on board the transit vehicle, mail the completed form back, or complete the same survey online. The samples collected were very large: 92,000 for the MNR survey and 34,000 for the CTA survey. For the MNR survey, every one of more than 700 inbound (toward New York) MNR trains (weekdays, Saturdays, and Sundays) was surveyed. For the CTA survey, a sample of the trips of every CTA bus route and subway line was surveyed.

Both surveys included questions asking whether the respondent was willing to be contacted for follow-up market research and collected contact information to reach them. These questions were designed to allow the agencies to build a readily available database of customers. Respondents could provide some or all of the following: mailing address, phone number, and e-mail address. A telephone sample collected in this way is not analogous to the RDD sample discussed earlier because this sample is a mixed sample of cell phones and land lines—that is, it is the telephone number where respondents are willing to be reached.

The surveys also asked respondents a set of demographic questions such as age, gender, household size, household vehicle ownership, and household income. These questions were in addition to questions about their travel on the transit system.

Survey Results

For this case study, the research question can be reframed slightly: is the sample of respondents willing to be contacted for future survey work by e-mail representative of the overall sample, or would surveying only the e-mail respondents lead to coverage error? The following analysis compares the e-mail sample and the telephone sample (i.e., those willing to be contacted by telephone), with the overall sample. Figure 3 shows the segments the survey samples can be grouped into based on responses to the market research and contact information questions. Table 4 shows the sample sizes for each of the survey groups for the CTA and MNR surveys. In all cases, the groups have in excess of 1,000 respondents.

Figures 4 and 5 compare the income distributions for the survey groups for the MNR and CTA surveys, respectively. The income distributions are markedly different between the overall sample collected on MNR trains and that collected on CTA subways and buses. For the MNR survey, 56% of the sample had an income greater than $100,000 per year; for the CTA survey, only 14% of the

TABLE 4 Percentage of Respondents Willing to Be Contacted for Future Market Research by Survey Location

	MNR		CTA	
	Count	% of Overall Sample	Count	% of Overall Sample
Phone only	5,407	6	5,281	16
E-mail only	3,744	4	1,353	4
E-mail and phone	12,225	13	5,809	17
Not available for future research	70,624	77	21,557	63
Overall sample	92,000	100	34,000	100
Phone sample (phone only + e-mail and phone)	17,632	19	11,090	33
E-mail sample (e-mail only + e-mail and phone)	15,969	17	7,162	21

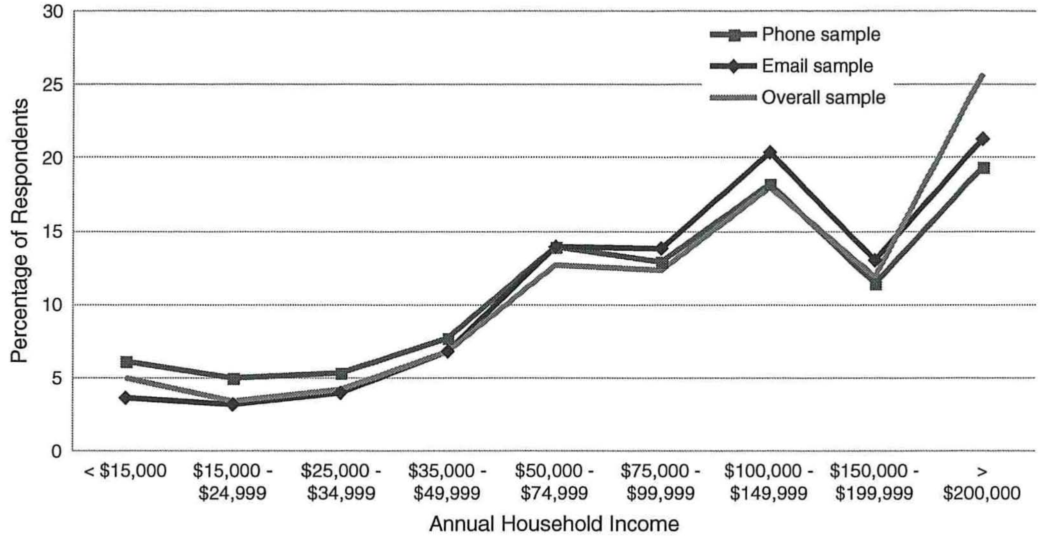


FIGURE 4 Household income by survey segment for MNR survey.

sample had that level of income. Most CTA riders had lower incomes; 62% had an income below $50,000 per year, compared with 19% of MNR riders.

Despite the differences between the overall sample income distributions, within the samples for each survey there are only small differences between the overall sample and the samples willing to be contacted by telephone or e-mail for future market research. The largest deviations between the samples occur between the group with income greater than $200,000 in the MNR survey and the group with income less than $15,000 in the CTA survey. For the MNR survey, both the telephone and e-mail samples underrepresent the proportion of respondents with incomes greater than $200,000, possibly due to reluctance among that group to share contact details on a survey form. For the CTA survey, the phone sample overrepresents the proportion of respondents with incomes under $15,000 and the e-mail sample underrepresents their proportion. This problem may be due to lower access to e-mail in this lower income group and less resistance to sharing a telephone number on a survey form.

Table 5 compares the age distributions for the survey groups. As with the income distribution, there are some differences between the overall samples collected on the MNR and CTA systems. The age of the MNR sample skews slightly older than the CTA sample. The within-sample comparisons also demonstrate only very small deviations between the telephone or e-mail sample and the overall sample.

Case Study 2 Conclusions

The results presented in this case study are based on surveys that took place on two large transit systems and obtained very large samples with high proportions of total ridership surveyed. Thus, there is a high level of confidence that the overall sample collected in each case is representative of the transit-riding populations that use those systems. The systems have very different demographic profiles, particularly for the income dimension: MNR riders have higher incomes

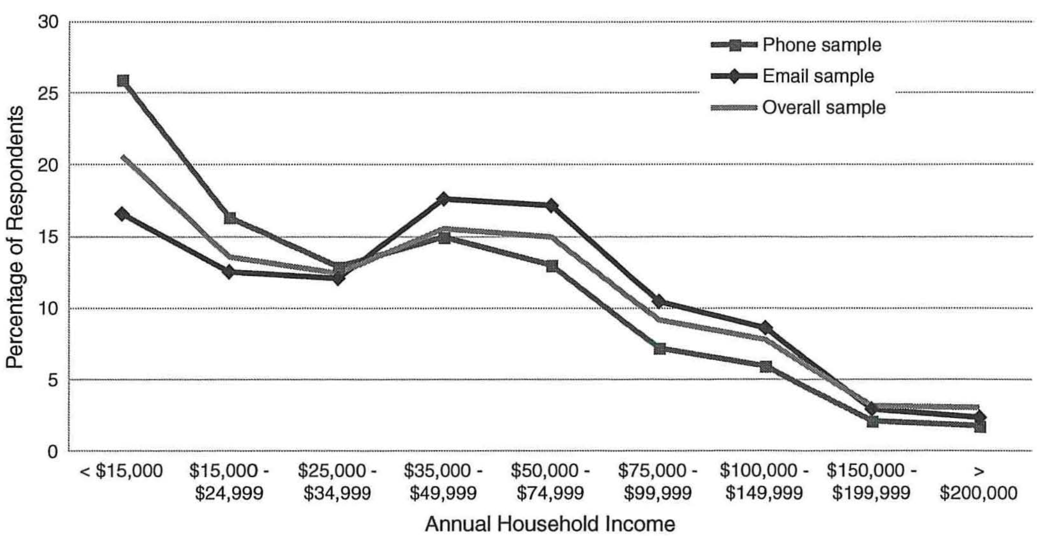


FIGURE 5 Household income by survey segment for CTA survey.

TABLE 5 Age by Survey Segment for MNR and CTA Surveys

Age (years)	MNR Phone Sample (%)	MNR E-Mail Sample (%)	MNR Overall Sample (%)	CTA Phone Sample (%)	CTA E-Mail Sample (%)	CTA Overall Sample (%)
Younger than 18	2	2	2	4	5	5
18–24	11	11	13	16	19	20
25–34	19	21	22	25	32	28
35–44	24	25	23	21	19	18
45–54	25	25	22	21	15	17
55–64	14	13	12	10	8	9
65 and older	5	4	5	4	2	4
Total	100	100	100	100	100	100

(25% are >$200,000), while CTA riders have relatively low incomes (35% are <$25,000).

For both the MNR and the CTA samples, the results show that an e-mail or telephone sample that could be drawn from the overall sample would be representative, at least along the income and age dimensions that have been shown to be important factors for telephone and Internet access and use. The results suggest that using an e-mail or telephone contact list for future market research from surveyed transit riders would not lead to significant coverage error in those future surveys.

The results indicate small differences between the e-mail and telephone contact lists and the full sample of riders, some of which were found to be statistically significant differences when applying a χ^2 test or a comparison of a two samples t-test. However, these findings are due to extremely low sampling errors around the distributions because of the very large sample sizes. A more reasonable approach to deciding whether the e-mail or phone distributions are close enough to the overall distributions to be considered similar is by judgment and inspection rather than by statistical tests. In the authors' opinion the e-mail and phone distributions are similar enough to the overall distribution to be considered representative of the overall population.

To mitigate any coverage error that is present, approaches such as weighting the data or additional data collection on the transit vehicle to correct for undersampling of certain demographic groups should be considered.

CONCLUSIONS

The research question posed at the beginning of this paper was "is coverage error introduced if only respondents who have Internet access are recruited to complete a travel survey?" The two case studies presented the following situations:

1. Potential road users intercepted at activity locations such as shopping malls and
2. Transit riders surveyed on the transit vehicle.

In both cases, the sample of survey respondents with Internet access or willing to be surveyed by e-mail was not substantially different from the overall sample along dimensions such as household income, age, and vehicle ownership. Therefore, it is possible to conclude that, for surveys of general populations of drivers or transit riders, surveying only those with Internet access does not introduce significant coverage error into travel survey samples. It is suggested that weighting data along demographic categories or conducting some supplemental intercept sampling will mitigate the small amount of coverage error that may be present.

The findings presented here are important for practical applications because collecting data with online surveys can be relatively inexpensive compared with collecting data using telephone surveys (or collecting data at activity locations or on transit vehicles). The finding that coverage error in Internet samples is not worse than, and is potentially smaller than, the coverage error in samples collected in telephone surveys means that collecting data with online surveys is a suitable alternative to conducting telephone surveys.

Additional research planned by the authors will extend the work presented here by comparing travel behavior among populations with and without Internet access. If travel behavior is found to be similar in those with and without Internet access, then the need to collect a supplemental intercept sample in addition to using online surveys would be reduced.

REFERENCES

1. Dillman, D. A. *Mail and Internet Surveys: The Tailored Design Method,* 2nd ed. John Wiley and Sons, Inc., Hoboken, N.J., 2007.
2. Tucker, C., J. M. Brick, and B. Meekins. Household Telephone Service and Usage Patterns in the United States in 2004: Implications for Telephone Samples. *Public Opinion Quarterly,* Vol. 71, No. 1, 2007, pp. 3–22.
3. Blumberg, S. J., J. V. Luke, G. Davidson, M. E. Davern, T. Yu, and K. Soderberg. Wireless Substitution: State-Level Estimates from the National Health Interview Survey, January–December 2007. *National Health Statistics Reports, No. 14.* National Center for Health Statistics, Hyattsville, Md., 2009.
4. Blumberg, S. J., and J. V. Luke. Coverage Bias in Traditional Telephone Surveys of Low-Income and Young Adults. *Public Opinion Quarterly,* Vol. 71, No. 5, 2007, pp. 734–749.
5. Link, M. W., M. P. Battaglia, M. R. Frankel, L. Osborn, and A. H. Mokdad. Reaching the US Cell Phone Generation: Comparison of Cell Phone Survey Results with an Ongoing Landline Telephone Survey. *Public Opinion Quarterly,* Vol. 71, No. 5, 2007, pp. 814–839.
6. Couper, M. P., and P. V. Miller. Web Survey Methods: Introduction. *Public Opinion Quarterly,* Vol. 72, No. 5, 2008, pp. 831–835.
7. *Pew Internet and American Life Project Surveys: Internet Adoption Chart.* Pew Research Center, Washington, D.C., 2010. www.pewinternet.org/Static-Pages/Trend-Data/Internet-Adoption.aspx. Accessed Nov. 15, 2009.
8. *Pew Internet and American Life Project Surveys: Demographics of Internet Users.* Pew Research Center, Washington, D.C., 2010. www.pewinternet.org/Static-Pages/Trend-Data/Whos-Online.aspx. Accessed Nov. 15, 2009.

The Transportation Planning Applications Committee peer-reviewed this paper.

Corridor Approaches to Integrating Transportation and Land Use

Kathleen Rooney, Karen Savage, Harrison Rue, Gary Toth, and Marie Venner

Transportation agencies nationwide are under pressure to help address a wider range of transportation issues than ever before in the United States. Many of them extend beyond traditional state department of transportation (DOT) activities and span of control. There is an interconnection between transportation and land use that the public and local decision makers do not often see. Yet the public frequently holds DOTs responsible for solving transportation problems resulting from local and regional land use decisions and preferred development patterns. The objective of this research is to identify and explore successful innovations in integrating transportation and land use planning for transportation corridors, with a focus on practices that could be transferred to other locations. A case study approach was used to identify projects that integrated, rather than merely linked, land use and transportation planning and decision making. This paper summarizes six case studies: Chicago Metropolitan Agency for Planning, Illinois; Envision Utah and the Mountain View Environmental Impact Statement, Utah; Gateway Route 1, Maine; NJFIT: Future in Transportation, New Jersey; UnJAM 2025 and Places29, Virginia; and MetroVision and Blueprint Denver, Colorado. The paper also analyzes practices and lessons learned, highlighting common themes among the case studies.

Transportation agencies nationwide are under pressure to help address a wider range of transportation issues than ever before in the United States. From creating new requirements to increasing awareness of climate change impacts on and by transportation, there is a great need for interagency collaboration to address shared issues efficiently and effectively. Furthermore, ambitious integration efforts must occur within the constraints of declining revenues, erosion of the purchasing power of state department of transportation (DOT) funds, and increasing public and political scrutiny. Many of these issues extend beyond traditional state DOT activities and span of control.

An interconnection exists between transportation and land use that the public and local decision makers do not often recognize. Yet, they frequently hold DOTs responsible for solving transportation problems resulting from local and regional land use decisions and preferred development patterns. The latter have presented a challenging context for DOTs to maintain the throughput capacity of their transportation system investments, while addressing local desire for more multimodal choices and improved access. Several states have made forays into integrating land use and transportation decisions and collaborating with local and regional agencies.

The objective of this research is to identify and explore successful innovations in integrating transportation and land use planning for transportation corridors, focusing on specific practices that could be transferred to other locations. After an extensive literature review, the research team focused on integrating land use and transportation planning and decision making, rather than on projects that only linked transportation and land use planning. For example, many state DOTs will engage a diverse set of stakeholders to collaborate on transportation solutions and offer their comments based on local land use, environmental resources, and community concerns. Sometimes localities align their comprehensive plans according to these transportation projects or enter into agreements to that effect. Although this endeavor is worthwhile, it only links land use and transportation rather than thoroughly integrating land use and transportation decisions. Each case study included one or more of the following elements:

- Innovative aspects or differing approaches that make the case study relevant for use by other DOTs;
- Simultaneous land use and transportation planning, with both sets of agencies at the table creating and exploring solutions together;
- A joint vision of desired outcomes across transportation, land use, and potentially environmental considerations;
- A strategic implementation plan or component with defined outcomes;
- Planned and actual resource allocation in infrastructure based on agreed upon desired outcomes; and
- Various catalysts for project initiation.

Integrated land use and transportation planning frequently can occur at a higher level than corridor planning, such as at the subarea or regional planning levels, with the results of the integrated decision making carrying through to specific corridor plans. Many state DOTs acknowledge that "corridors" and their transportation impacts do not end at the right-of-way boundaries; to be successful in integrating transportation and land use, corridor studies must extend beyond the right-of-way. As a result, notable practices may not reflect the traditional definition of a transportation corridor, although they lay the groundwork for corridor and project-level integration. With this broader definition, the research team identified case studies representing a more expansive framework where integration is planned in addition to those specific corridors where integration has been implemented.

K. Rooney, ICF International, 1725 I Street NW, Suite 1000, Washington, DC 20006. K. Savage, ICF International, 9300 Lee Highway, Fairfax, VA 22031. H. Rue, ICF International, 2222 Chapel Hill–Nelson Highway, Suite 480, Durham, NC 27713. G. Toth, Gary Toth Consulting LLC, 19 Cottage Street, Lambertville, NJ 08530. M. Venner, Venner Consulting, 9947 West Oregon Place, Lakewood, CO 80232. Corresponding author: K. Rooney, krooney@icfi.com.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 42–49.
DOI: 10.3141/2176-05

CASE STUDY SUMMARIES

Six case studies were examined: Chicago Metropolitan Agency for Planning, Illinois; Envision Utah and the Mountain View Environmental Impact Statement (EIS), Utah; Gateway Route 1, Maine; NJFIT: Future in Transportation, New Jersey; UnJAM 2025 and Places29, Virginia; and MetroVision and Blueprint Denver, Colorado. For each case study, the team completed background research and conducted a minimum of four interviews with relevant participants and experts representing different viewpoints.

Chicago Metro Agency for Planning

In 2005, the existing regional planning organization, Northeastern Illinois Planning Commission, and the transportation planning agency, Chicago Area Transportation Study, were merged at the direction of the state legislature. The new comprehensive agency became known as the Chicago Metro Agency for Planning (CMAP). This consolidation was enacted to accomplish the region's aim to integrate land use and transportation. Regional leaders concluded that it was needed to accommodate the additional 2.8 million residents and 1.8 million jobs anticipated for the Metropolitan Chicago Region by 2040 (*1*).

From the beginning, CMAP recognized that a regional vision is essential to guide future growth and is the foundation for the Chicago area's first truly regional comprehensive plan, GoTo2040. The plan will go beyond traditional land use and transportation performance measures and incorporate broader goals such as health and human services. Scheduled for completion in fall 2010, it will establish policies and strategies to guide growth and development in the region through 2040. Notable practices include the following:

- Creating an integrated agency and breaking down the barriers between the former ones;
- Making the case for a regional vision and getting the best input possible;
- Encouraging full, active, and collaborative involvement of individual municipalities to align and input municipal plans and needs into the regional vision;
- Leveraging Chicago's intellectual capital; and
- Assessing developments of regional importance.

Envision Utah and the Mountain View Environmental Impact Statement

Initiated by business and civic leaders, Envision Utah is well known as an excellent example of land use and growth visioning. It was one of the first attempts in the country to conceptualize a statewide future to include such a large number of diverse stakeholders. In addition to the statewide visioning process, Envision Utah provides on-site technical assistance for communities to implement the statewide vision.

Envision Utah is involved in a variety of local implementation projects. One of them involves the Mountain View corridor southwest of Salt Lake City. Envision Utah staff, along with the Utah Department of Transportation (UDOT), facilitated the decision-making component of the Mountain View EIS (also referred to as Growth Choices) to integrate this transportation project and local land use effectively and cohesively. According to UDOT, this project provided an opportunity to integrate transportation and land use at the beginning of a project and avoid some of the traditional transportation EIS conflicts (*2*). Notable practices include the following:

- Connecting statewide visioning to local corridor implementation,
- Using upfront agreements to keep everyone committed and at the table throughout the process, and
- Including a variety of options as part of a balanced transportation solution, not as mitigation or appeasement.

Gateway 1

Midcoast Maine residents were looking for a more collaborative approach to address regional concerns about increasing traffic congestion and truck traffic due to rapid corridor development along U.S. Route 1. Conflicting opinions about the corridor became evident during the state-sponsored Regional Transportation Advisory Committee processes, which solicited advice from communities on how Maine should invest its transportation dollars (K. Fuller, personal communication, Nov. 14, 2008). The Maine DOT therefore needed to find a proactive way to work with, and be responsive to, the 21 communities in the Midcoast region. The Midcoast region committee suggested that the Maine DOT develop a comprehensive plan for the corridor instead of reacting to "spot" problems, such as highway widening. In response, the Maine DOT initiated the Gateway 1 process—a long-term strategic planning project that sought to find a way to integrate community involvement and combine municipally based land use and state-based transportation planning. The Maine DOT led a community-based initiative engaging the residents of the 21 communities in the corridor. They held more than 50 community and larger regional meetings where information such as a visual assessment of the corridor and a basic land use, transportation, environmental, and community inventory was presented (K. Fuller, personal communication, Nov. 14, 2008). From these meetings, the Gateway 1 Stakeholder Steering Committee was created and tasked with developing the corridor vision and drafting the implementation plan. The goal for this initiative is to preserve mobility while enhancing safety, transportation choice, economic strength, and quality of life along the corridor. Notable practices of Gateway 1 include the following:

- Fostering collaboration among the project team instead of using the team to sell DOT ideas,
- Creating lasting institutional arrangements so that commitment to long-term solutions remains,
- Balancing transportation needs and the community's concerns,
- Using strategies to equalize economic development benefits among communities in the corridor,
- Tailoring public outreach to the community and trying new techniques to get each community interested, and
- Leveraging supportive state legislation to improve incentives for change.

New Jersey Future in Transportation Program and New Jersey Routes 9, 29, 57, 31–202, and 322

The state of New Jersey has faced serious congestion problems over the last two decades. The New Jersey DOT has relied on the standard approach of adding capacity and road widening but has found it to be an unsustainable solution to congestion. Financial resources have declined and increasingly need to be used to maintain and fix the

existing infrastructure. Due to community resistance and extended project development periods, it has taken decades to deliver some capacity expansion projects. Consequently, the New Jersey DOT decided to develop an approach that addressed congestion problems with a balanced and long-term focus rather than chasing short-term but elusive fixes. The initiative, known as NJFIT: Future in Transportation (NJFIT), focuses on integrating land use and transportation planning within the context of viable regional corridor projects rather than simply creating short-term solutions. The strategy has proven successful by encouraging municipalities to think beyond transportation improvements to develop sustainable land use policies that complement and support transportation strategies. Notable practices used by the NJFIT program include the following:

- Helping communities to understand how codes, zoning, and other ordinances can steer development into unsustainable patterns;
- Providing technical assistance and toolkits to help communities create a codified, shared vision for the community;
- Establishing a statewide focus on integrated land use and transportation to promote future project successes strategically; and
- Finding the appropriate land use, access management, and local network solutions to supplement and reinforce the traditional DOT approach of investing in the state highway system.

UnJAM 2025 and Places29, Virginia

The United Jefferson Area Mobility Plan (UnJAM 2025) combines the metropolitan planning organization (MPO) long-range transportation plan with a rural area plan for five counties surrounding Charlottesville, Virginia. Staffed by the metropolitan planning agency, Virginia DOT, local land planners, and transit agencies, UnJAM 2025 integrated transportation planning with local comprehensive land use plans. It called for a corridor-based approach to reengineering and investing in existing roadways while coordinating developer investments to produce a multimodal network of local streets, which will protect the capacity of major thoroughfares while increasing multimodal travel choice. Places29 is the corridor implementation of the UnJAM 2025 principles. The synchronization of the processes (decision making and schedule) established a vision to accommodate growth along U.S. Route 29 in a centers-based approach. It involved coordinating transit-ready developer investments with a multimodal network of new streets for local travel, an expanded regional transit system, urban grade-separated intersections at key locations on the route, and a detailed access management and incident management plan (*3*). Notable practices of this case study include the following:

- Cooperating at multiple levels of government, including the MPO, relevant counties and cities, and state agencies;
- Creating an incremental and synchronized solution toward a common vision in Places29; and
- Coordinating the land use and transportation processes seamlessly through one public involvement process and brand.

MetroVision and Blueprint Denver

Denver's MetroVision Plan is a regional plan for growth and development, which incorporates transportation, land use, and environmental concerns into a long-range regional plan to manage growth and encourage more efficient and effective transportation investment and land use. The product is not just a compilation of local plans but a shared vision for the region with targeted growth areas agreed upon by local metro area municipalities. Blueprint Denver is an example of a local plan that has grown out of the regional framework set by the Denver MetroVision Plan. It highlights specific steps to support urban centers, environmental quality, and a balanced, multimodal transportation system. Notable practices implemented in developing the plans are as follows:

- Using alternative initiatives and tools, such as the Living (Complete) Streets approach, a new street classification system, and appropriate upzoning to make corridor and regional visions a reality;
- Investing in the most accurate, up-to-date tools and technologies that support the prioritization of and lay the groundwork for the most pertinent decisions for the situation at hand; and
- Ensuring that a decentralized area commits to a shared vision through formal agreements and continued education by using vertical and horizontal collaboration.

LESSONS LEARNED AND NOTABLE PRACTICES

Although each case study is extremely valuable, certain themes and practices among all six are just as important. In this section, lessons learned and notable practices from all case studies are discussed.

Lessons Learned

A comprehensive analysis of the different practices in each case study yielded several overarching themes, which are summarized below. More detail is provided in the individual case studies.

Local Communities Will Not Always Resist State DOT Involvement in Land Use Planning

Local communities can and will welcome state involvement in land use planning. Cooperation depends on the approach of the state DOT and how the DOT addresses local concerns. Success or failure in collaborating with communities on land use planning depends on how the community perceives the state's involvement, attitude, and objectives. In Utah, the state brought in Envision Utah to help create land use and transportation scenarios, building on the broad base of support and credibility of Envision Utah (*2*). In Maine, state and local communities along the state route are full partners in crafting a plan for the corridor. The DOT spent the necessary time up front to build trust and good working relationships, which have been essential in developing a plan supported by the communities and the state DOT. In Virginia, the state played a leading role in engaging communities and other agencies to cooperate and contribute to the planning process. The key is for the state to approach communities with transparent goals and to demonstrate a willingness and ability to listen to local concerns. A key approach is to work through existing regional entities where the localities and state agencies have established relationships.

Land Use and Transportation Planning and Decision-Making Activities Occur Whether DOTs or Other Agencies Seek to Integrate Them

The presumption that transportation agencies cannot influence land use is false. Integration can occur by collaborating with communities

or through direct and indirect impacts on land use arising from transportation investment decisions. While transportation agencies may understand that land use decisions are local, it does not mean their transportation systems are not coordinated with land use or that there are no land use impacts from transportation decisions. State elected officials, MPOs, and DOTs alike often hesitate to discuss land use, fearing that localities may react poorly to perceived state interference in this local domain. However, localities can, and in many cases do, welcome state and regional leadership and facilitation. At the same time, local entities do not want outsiders to take control of the local planning process and force specific or top-down solutions.

While this lesson learned was specific to the Gateway Route 1 initiative, it was exemplified in most of the other case studies—typified as either the fear of transportation agencies becoming involved in land use decisions or as an attempt at a solution to address this fear. The Gateway initiative shows that the manner of participation by state agencies is a key determinant in creating and implementing an integrated transportation and land use solution. The Gateway initiative demonstrates that even in a strong home-rule state, if the DOT or other state agencies approach communities with respect for local knowledge and authority and a spirit of collaboration, those communities are responsive to outside influence aimed at making integrated transportation and land use solutions work for specific challenges.

Collaborative and Integrated Solutions Are Achievable, Especially If Agencies Start with a Blank Slate

Many case study interviewees mentioned that if the stakeholders involved in the planning process set aside their preconceived ideas about what the "correct" solution may be before entering the process, then new and perhaps better solutions were often possible. CMAP began with an entirely new mandate of integrating land use and transportation and other important issues, while integrating actions within the agency as an institution. In the Growth Choices process in Utah, a key stakeholder cited that starting with a blank slate was the only way to see the linkages between the impacts of transportation and land use decisions (*2*). In the Gateway Route 1 initiative, Maine DOT refused to allow discussion of possible solutions until everyone agreed on the problems in the corridor (*2*). For the oldest projects analyzed for this research (e.g., the NJFIT example), some of the integrated solutions have been fully implemented. Similarly, with UnJAM 2025 and Places29 in Virginia, the DOT has begun to implement some of the initial recommendations of the transportation plans and has continued to fund and support similar integrated planning throughout the state (*2*). In Denver, this integration has been codified through the street classification system, which identifies land use and transportation standards for different types of streets. Each of these case studies illustrates that integrating transportation and land use can be successful if fresh perspectives are incorporated with innovative approaches.

Integration of Land Use and Transportation Can Happen Within Any Organizational Structure at Any Level and Apply to Wide Variety of Transportation Contexts

The wide range of case studies presented illustrate that integrating transportation and land use can occur at different levels and in different contexts. The case studies ranged across the country from the Northeast (Maine) to the mountain states (Colorado) and Southeast (Virginia). They also covered a wide variety of development contexts from urban (Chicago and Denver) to rural (Maine), or all the above (Utah, New Jersey, and Virginia). They spanned different levels of agency initiatives. CMAP is an MPO-level agency in an urban area that will be integrating land use and transportation in every project and for all regional issues. In Utah, the Mountain View EIS was initiated at a state level and applied to a very traditional transportation agency activity (EIS). However, integrated land use and transportation scenarios were developed and evaluated, and the solutions will be implemented at local and state levels. Although Gateway Route 1 began as a state DOT-led initiative, it is evolving into a regional corridor-based framework (*4*). NJFIT is an overarching state DOT program with consistent staff statewide. UnJAM 2025–Places29 is also an MPO-led initiative but in a much smaller community 2 h outside of Washington, D.C. In Denver, while it was an MPO vision, much of the implementation innovation is occurring within the city of Denver. The research demonstrated that it is not necessary to have enabling or special legislation or a unique situation to succeed in integrating land use and transportation planning and decision making.

State Legislation Can Provide Structure Needed to Support Land Use and Transportation Integration

The most obvious example of how state legislation can support the integration process is illustrated by the CMAP process in Chicago where state legislation mandated integrating the Northeastern Illinois Planning Commission and the Chicago Area Transportation Study. Another example is the Sensible Transportation Policy Act (STPA) passed by the Maine legislature, which directed Maine DOT and the state planning office to draft a rule to link land use and transportation processes of the STPA to those for the Comprehensive Planning and Land Use Regulation Act. The legislation codified the premise that land use and transportation planning must be done simultaneously to protect transportation safety and mobility while enhancing communities. Another example is in New Jersey, where the pilot corridor projects involved the New Jersey DOT and the New Jersey Office of Smart Growth and frequently the New Jersey Office of Environmental Protection and municipal compacts or partnership agreements.

Giving Travelers More Options Is One of Most Common Solutions

The principles of a traveler-based strategy include designing for pedestrians and bicyclists; creating more compact, mixed-use downtowns with connected street networks; connecting transportation modes, particularly around transit; and considering that congestion in dense city centers can be effective at reducing automobile travel. These strategies support multimodalism, ultimately reducing traffic congestion and creating a sense of place for the community. Additional toolbox strategies include the New Jersey DOT's building for transit, creating more connections, providing better access, designing roads in context with their surroundings, and calming traffic. Such solutions also tend to help DOTs control access to their facilities and preserve past and current investments in the highway system. In Virginia, they tried to create more complete, connected road and parallel road networks on a grid system and include multimodal design aspects, coordinated with future transit-ready development (Figure 1).

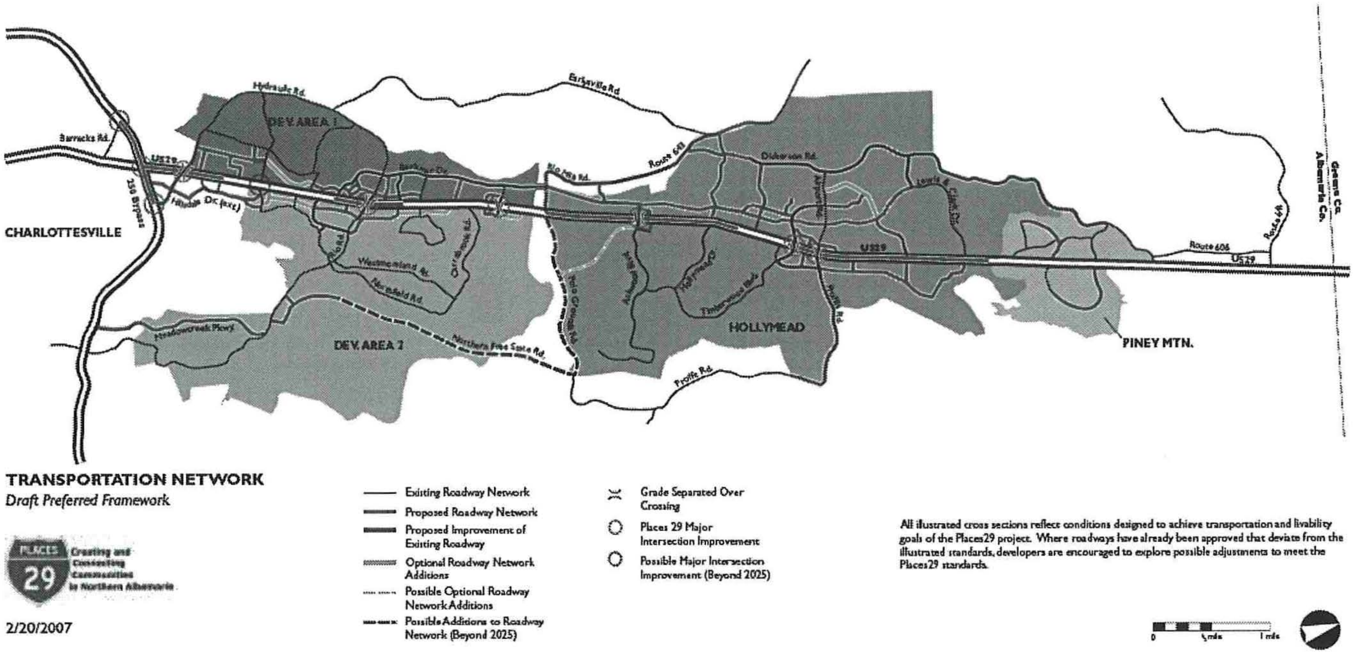


FIGURE 1 Places29 identified how a multimodal road network parallel to typical suburban highway could be provided over time by coordinating and connecting roadways built by individual developers. (Source: H. Rue, January 6, 2010.)

Any Worthwhile Process Takes Time but Can Be Incredibly Valuable over Long Term

Because of limited resources, communities are often used to competing with each other and may not perceive much incentive to cooperate in achieving potentially shared goals. However, in Maine, New Jersey, Virginia, and Denver, the sponsoring transportation agencies took the time to earn the community's trust and build consensus. This investment led to mutual understanding and shared decision making among community members. Sometimes it may appear that a year or more for consensus building is a luxury. However, patient, upfront collaboration with key stakeholders can lead to eventual widespread support for the right project in the right place and to cost savings for the design and construction of projects that are supported by the community. Regions that have an established agency (typically, a regional planning agency or council of governments) that incorporates MPO and rural transportation planning, rideshare and transit system coordination, housing programs, and land use planning will have a head start on an integrated approach—even when land use decision making is retained by the individual localities.

Transportation Agencies Have Tools Necessary to Succeed at Integrating Transportation and Land Use

In addition to identifying the most notable practices, the project team also explored the breadth of solutions used. As shown in Table 1, each case study project used a variety of practices to achieve integrated solutions. Several strategies are used in all case studies (or almost all) and are worth mentioning. They include the following:

- Focusing on access management (in design and operations);
- Integrating land use and facility design to address capacity, aesthetics, safety, and multimodal issues;
- Providing for multimodal options in design;
- Improving transit;
- Rezoning for increased density;
- Regulating development to match corridor form; and
- Controlling land uses adjacent to transportation facilities.

Agencies do not necessarily need to develop new tools but can use the tools at hand to integrate land use and transportation at the corridor level. While it might be easier from an organizational point of view to integrate land use and transportation planning fully in one agency, as in the Chicago case, it is possible to succeed in integrating land use and transportation planning and decision making under a variety of organizational structures.

Notable Practices

Some of the most valuable and noteworthy practices exemplified by the six case studies in the research are summarized in this section.

DOT Funding of Local Studies for Integrated Planning or Smart Growth

In developing the NJFIT program, planning was done in full collaboration with the communities, which were allowed to direct consultant resources in creating a vision for how each community wanted to evolve. With NJFIT, each corridor project proceeded through three identifiable stages: public education and outreach to elevate public awareness and point out the difficulties of pursuing "business as usual;" community consensus building through use of external consultants, information sharing, and shaping of shared values; and community codification that involved revising municipal plans and ordinances and drafting agreements that involved multiple jurisdictions (P. Cohn and G. Toth, unpublished data, Dec. 18, 2009). This

TABLE 1 Spectrum of Solutions

	CMAP, Ill.	Envision Utah–Mountain View EIS, Utah	Gateway Route 1, Maine	NJFIT	UnJAM 2025–Places29, Va.	MetroVision–Blueprint Denver, Co.
Design						
Rightsizing the road		P			P, I	
Access management	P	P	P	P, I	P, I	P
Integrating land use and facility design to address capacity, aesthetics, safety, and multimodal issues		P	P, I	P	P, I	P, I
Aesthetic improvements to better integrate transportation facilities			P	P	P	P, I
Provision for multimodal options	P, I	P	P	P	P, I	P, I
Covering–depressing of roads to reconnect neighborhoods					P	
Creating redundancy–parallel roads in the network		P	P	P	P, I	
Operations						
Intelligent transportation strategies	P			I	P, I	P
Variable-priced managed lanes	P	C				
Overall freeway land management		P				
Access management	P	P	P	P, I	P	
Services and Programs						
Transit improvements	P	P	P		P, I	P
Commute trip reduction programs	P				P, I	
Park & Ride lots linked to high-occupancy vehicle lanes	P				C	
Land Use						
Rezoning to get transit-oriented development, higher densities, and mixed-use clustering		P, I	P, I	P, I	P, I	P, I
Development regulation to match corridor form		P, I	P, I	P	P, I	P, I
Growth management		P, I	P, I		P	P, I
Concurrency requirements	P		P, I			P, I
Joint use of Park & Ride lots and development		P			P, I	
Controlling land uses adjacent to transportation facilities		P, I	P, I	P	P, I	P, I
Protecting adjacent land use from undesirable aspects of transportation facilities		P, I	P, I		P	P, I
Other						
Transfer of development rights programs			P	P	C	
Intergovernmental agreements for reciprocal action		P, I	P, I	P	P, I	I

NOTE: P = planned; I = implemented; C = considered but not implemented.

practice may be contrasted with some other DOTs' corridor planning approaches in long-range planning, in which the DOT directs the consultant resources, and regional planning agencies are asked to approve results. Similarly, in Virginia and Utah, state DOTs contributed to funding their integrated planning initiatives (*2*).

Educating Communities About Trade-Offs and Implications of Different Investment Strategies, Including Consideration of Megatrends, and About Land Use and Transportation Concerns

In New Jersey, the NJFIT program is helping communities understand how codes, zoning, and other ordinances can steer development into unsustainable patterns. The program also provides technical assistance and toolkits to help communities create a codified, shared vision for the community (*5*). Denver and Chicago allowed advisory boards and the public to share values and explore investment scenarios and implications to come to a consensus and understand more about trade-offs. The Utah DOT evaluated land use and transportation in integrated scenarios, ranging from a low density, automobile-dependent scenario to one with compact, dense development and an extensive transit component. The Maine DOT followed a similar approach by developing and considering the impacts of three integrated scenarios, which the stakeholder steering committee evaluated. DOTs, MPOs, and corridor planning groups can draft "myth busters" to help the public understand land use and transportation connections, as the Maine DOT Gateway Route 1 Corridor planning group did to build more support for the Gateway Route 1 vision and implementation (*2*). NJFIT and Denver's Blueprint program offer planning resources on their websites that address some of these issues as well. UnJAM 2025 used interactive tools at the planning sessions to engage citizens constructively.

Relating Mobility to Community Form

Denver's Living Streets program, street classification, and rezoning efforts have helped relate mobility to the community's form and design. The New Jersey DOT initiated a Mobility and Community Form (MCF) program to help communities plan future transportation and land use. MCF planning emphasizes the connections between the local system and the design of community facilities, buildings, and open space. This program is supported by the New Jersey DOT and the Municipal Land Use Center at the College of New Jersey (*6*). The program helps communities transition from traditional zoning in master plans to a more integrated form-based development code that links land use and transportation. The approach encourages linking the local street grid to the design of community facilities, buildings, and open space. The New Jersey DOT developed a guidance resource for communities to use to incorporate these strategies into their master plans.

Focusing on "System Wellness" and Regional Approach That Identifies Challenges for Overall Transportation Network

NJFIT is designed to focus new investment on keeping the transportation system healthy rather than waiting for it to deteriorate and then doing the infrastructure equivalent of major surgery. This focus on "system wellness" means more smaller projects that are synchronized with county and local transportation systems and land use plans that can be implemented faster through state, regional, and local partnerships. It is a "faster, better, cheaper approach to new capital investment" (H. Volk, personal communication, Nov. 20, 2008).

In the Denver metro area, development of the Mile High Compact is a landmark intergovernmental agreement documenting and fostering a broad base of support for growth management in the region. By signing on to the agreement, 40 jurisdictions representing more than 80% of the Denver region's population voluntarily agreed to designate and abide by a voluntary urban growth boundary or area, to accept their share of future land development, and to identify common comprehensive plan elements and MetroVision objectives in their local comprehensive or master plans.

In Chicago, CMAP created a process to review major developments that have significant regional effects. Developments qualify if they meet or exceed certain thresholds related to size, purpose, and intensity of use. While localities maintain final approval authority, CMAP's role is critical to identify broader potential benefits and drawbacks.

Building Institutional Arrangements and Processes for Decision Making That Allow Regional Residents to Address Transportation, Land Use, and Environmental Issues Collaboratively

In Maine, this forum has already led to smaller voluntary localized guidelines that support integrated planning goals and help to achieve corridor objectives. Through a memorandum of understanding in the alternative institutional arrangements, localities are codifying agreements that help to implement a transit-oriented corridor through local laws, regulations, and comprehensive plans (*4*). In Chicago, CMAP was created on this premise by the state legislature.

Explore Community Values in Depth to Build on Them and Find Consensus

In Maine, New Jersey, Virginia, and Denver, the sponsoring transportation agencies took the time to earn the community's trust and build consensus. This investment led to mutual understanding and shared decision making among community members. For example, the Maine DOT conducted an attitudes survey to help build the collaborative process and guide the vision and implementation of the Gateway 1 initiative. More than 500 randomly selected residents across the region were asked about specific values related to property rights, governmental regulation, home rule, interlocal cooperation, economic development, scenic quality of the corridor, and choice of transportation (*7*). This information helped create a Gateway 1 specific solution with messages tailored to each community. The Midcoast residents' values guided the solutions and, ultimately, their ability to be implemented.

DOTs and MPOs Have Created Incentives and Pursued State Legislation or Bond Initiatives to Implement Integrated Planning and Sustainability Programs

The Maine DOT, the New Jersey DOT, and to a certain extent Virginia have taken this approach. In Chicago, CMAP has pursued bond initiatives to create financial incentives for local governments to support desired development approaches. In Maine, municipalities that develop plans using the STPA guidelines are eligible for transportation-planning assistance and other investment incentives such as bonus prioritization points for the Maine DOT's competitive programs and funded highway reconstruction and mobility projects and incremental reductions in local match requirements. The New Jersey DOT emphasizes that state and federal transportation funding to implement transportation improvements is linked to the municipality's willingness to embrace integrated land use and transportation principles in master planning and zoning ordinances. The Virginia localities have pursued state legislation to allow development impact fees and to allow local funding options for transit improvements, but their efforts have been rejected by the state legislature (*2*).

DOTs Have Found Success in Pressing the Land Use Connection and a Quid pro Quo of Local Investment in Land Use Decisions That Support Multiple Modes and Help Preserve the DOTs' Highway Capacity Investments

NJDOT shares the message that "communities that want to have a say in transportation must let transportation agencies have a say in the community's land use." Maintaining this quid pro quo message upfront is an important tool to establish a message that compromise is essential (P. Cohn and G. Toth, unpublished data, Dec. 18, 2009). Completing a well-connected network of roadways parallel to major highways, with better connections within and between neighborhoods, was a key transportation principle advocated in the New Jersey example as well as in the Virginia case. Solutions in both states included well-executed design details for pedestrian-friendly streets, bike lanes and trails, transit stops, safer intersections, and pedestrian crossings.

CONCLUSION

Transportation agencies are increasingly being held responsible for addressing transportation challenges resulting from local and regional land use decisions and development patterns. Consequently, there is a great need for interagency collaboration to address shared issues at the nexus of transportation and land use efficiently and effectively. The objective of this research is to identify and explore successful innovations in integration of transportation and land use planning for transportation corridors, focusing on practices that could be transferred to other locations.

Rather than identifying projects that merely link transportation and land use planning, the research team identified six case studies from across the nation that exemplify successful strategies and approaches for integrating land use and transportation planning and decision making. This paper summarizes those approaches and analyzes notable practices and lessons learned, highlighting notable practices within each of the case studies and as a whole. Such practices demonstrate various approaches used by state DOTs and other agencies that have broadened the scope of planning and design of transportation facilities to improve management and function of regional transportation corridors, thereby better serving community needs. Through a comprehensive analysis of the different practices in each case study, the research team identified key lessons learned by organizations that are proactively integrating land use and transportation decision making, which include the following:

- Local communities will not always resist state DOT involvement in land use planning.
- Land use and transportation planning and decision-making activities occur whether DOTs or other agencies seek to integrate them.
- Collaborative and integrated solutions are achievable, especially if agencies start with a blank slate.
- The integration of land use and transportation can happen in any organizational structure and apply to a wide variety of transportation contexts.
- State legislation can provide the structure needed to support land use and transportation integration.
- Giving travelers more options is one of the most common solutions.
- Any worthwhile process will take time but is valuable over the long term.
- Transportation agencies have the tools necessary to succeed at integrating transportation and land use.
- Through this research, the team determined that successful integration often includes the following key elements: simultaneous land use and transportation planning, with both sets of agencies at the table creating and exploring solutions together; creating a joint vision of desired outcomes across transportation, land use, and potentially environmental considerations; having a strategic implementation plan or component with defined outcomes; and allocating resources and government investments in infrastructure based on agreed upon desired outcomes.

ACKNOWLEDGMENTS

This work was based on more detailed research conducted as part of NCHRP 8-36 Task 86. The authors thank the NCHRP 8-36 Task 86 panel and Lori Sundstrom as well as the following primary contributors who were interviewed as part of these case studies: Randy Blankenhorne of the Chicago Metro Agency for Planning, Alan Matheson of Envision Utah, Teri Newell of the Utah Department of Transportation, Kat Fuller of the Maine Department of Transportation, Carol Morris of Morris Communications, Paul Cohn of the New Jersey Department of Transportation, Judy Wiegand of the County of Albemarle, Harrison Rue of ICF International, Steve Gordon of the Denver Department of Community Planning and Development, Bill Johnston of the Denver Regional Council of Government, and Cynthia Patton of the Denver Department of Public Works. The authors also thank the team's research staff: Ryan Thompson, Liza Wallis, and Anna Chavis of ICF International. For more detailed information on each case study, see NCHRP 8-36 (Task 86) Corridor Approaches to Integrating Land Use and Transportation, available at http://144.171.11.40/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=2405.

REFERENCES

1. *Regional Vision Snapshot.* Chicago Metropolitan Agency for Planning, Chicago, Ill. www.cmap.illinois.gov/WorkArea/showcontent.aspx?id=1610. Accessed Nov. 24, 2008.
2. *Corridor Approaches to Integrating Transportation and Land Use.* ICF International, Washington, D.C., 2009. www.statewideplanning.org/_resources/252_08-36-86.pdf.
3. The Thomas Jefferson Planning District Commission, Charlottesville, Va. www.tjpdc.org/transportation/unjamp.asp. Accessed Dec. 9, 2008.
4. Gateway Maine website. www.gateway1.org/. Accessed March 5, 2010.
5. *Toolbox: Prescriptions for Wellness.* New Jersey Future in Transportation, Trenton. www.state.nj.us/transportation/works/njfit/toolbox/. Accessed Nov. 1, 2008.
6. *Mobility and Community Form.* New Jersey Department of Transportation, Trenton. www.state.nj.us/transportation/community/mobility/. Accessed Dec. 20, 2008.
7. *Executive Summary: Attitude Survey, Gateway 1 Corridor.* www.gateway1.org/new/phase1/Major_themes_to_explore_Apr_06.pdf. Accessed Oct. 15, 2008.

The Transportation Planning Applications Committee peer-reviewed this paper.

From Academia to Application

Results from Calibration and Validation of First Hybrid Accessibility-Based Model

Vincent L. Bernardin, Jr., and Michael Conger

This paper presents an overview of a new architecture for travel demand forecasting models. The new hybrid accessibility-based models are neither traditional trip-based models nor standard activity-based simulation models. Instead, they are a blend of aggregate and disaggregate component models offering much of the sensitivity typical of activity-based models but with more modest development and application costs. This paper presents results from the first full application of the new model design. It documents the experience of the Knoxville (Tennessee) Regional Transportation Planning Organization from the data collection and visioning stage to the stage of investing time and resources to develop the new model. Ultimately, some basic comparisons are made between the new hybrid model and the previous trip-based model. Some key innovations in the new model design are highlighted, making special note of added sensitivity to planning and policy variables. The paper also reports findings related to the effects of residence location on travelers' willingness to travel and the effect of psychological barriers such as river crossings and county lines on stop location choice.

TRB, in its *Special Report 288: Metropolitan Travel Forecasting, Current Practice and Future Direction,* encouraged metropolitan planning organizations (MPOs) to share their experiences developing advanced travel models, stating, "MPOs experimenting with or fully implementing advanced modeling practices should document their experiences, including costs, advantages, drawbacks, and any transferable data or model components." (*1*, p. 9). This paper reports on the development and implementation of an advanced travel model for the Knoxville (Tennessee) Regional Transportation Planning Organization (TPO). The first section presents the TPO's motivation and data collection efforts supporting the model development and the final section focuses on the development and application costs of the model and comparisons of model performance with the previous trip-based model where possible.

The remainder of the paper focuses on the model's architecture, highlighting innovations and the sensitivity offered to various policy and planning variables. The new Knoxville regional travel model (KRTM) is the first of a family of hybrid, accessibility-based travel models. It is neither a traditional trip-based model nor a standard activity-based microsimulation model. Instead, it is a blend of aggregate and disaggregate component models that offers much of the sensitivity typical of activity-based models with more modest development and application costs. The new model design was introduced in concept only 2 years ago (*2*). Since then various components of its design have been the subject of academic research (*3–5*). A second model of similar design is now under development for the Evansville, Indiana, MPO. This paper, however, presents the first application of the new model design for the Knoxville TPO and its sensitivity and performance.

V. L. Bernardin, Jr., Bernardin, Lochmueller & Associates, Inc., 6200 Vogel Road, Evansville, IN 47715. M. Conger, Knoxville Regional Transportation Planning Organization, 400 Main Street, Suite 403, Knoxville, TN 37902. Corresponding author: V. L. Bernardin, VBernardin2@BLAinc.com.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 50–58.
DOI: 10.3141/2176-06

DATA COLLECTION AND MODEL DEVELOPMENT PROGRAM

The development of any travel model should be supported by a good program of local data collection and guided by the planning needs and vision of a particular community. The development of the new KRTM could not have taken place without the TPO's data collection efforts; the model's architecture evolved in large part in response to events that identified improved model accuracy and sensitivity as priorities while acknowledging the constraints of the TPO's schedule and budget, which precluded development of a standard activity-based model.

Good, ongoing data collection is essential to developing and maintaining good models. It is worth chronicling the Knoxville TPO's data collection and model development program over the past decade because it illustrates how an ongoing data collection program can provide better data and be more financially feasible by spreading the cost of data collection efforts over several years. The history of development of the Knoxville model also illustrates good interagency cooperation between the Tennessee Department of Transportation (TDOT) and the Knoxville TPO with the helpful involvement of the University of Tennessee and a peer review panel.

In the year 2000, the TPO commissioned the 2000 Household Travel Behavior Study, a prompted recall survey of 1,538 households in Knox and Blount Counties to support the development of a modern four-step travel demand model to replace the existing travel model for Knox and Blount Counties implemented in MINUTP (*6*). The TPO contracted with a consultant to develop a model incorporating an expanded study area. The model area was expanded to incorporate neighboring counties using year 2000 census data and rich roadway attribute information from TDOT's Tennessee Roadway Information Management System. Model development began in 2003 and was completed in 2004, validated to year 2000 traffic counts.

In 2005, the TPO conducted a peer review of its travel model (*7*). The panel included representatives of several MPOs, the University

of Tennessee, FHWA, the TPO, and its consultant. The peer review made several recommendations:

- Additional data collection:
  - Transit onboard survey,
  - External cordon line origin–destination survey,
  - National Household Travel Survey add-on and household survey update including collar counties, and
  - Commercial vehicle survey.
- Further model development and refinements:
  - Develop mode choice models,
  - Develop a land use model,
  - Enhance network detail and zone structure,
  - Make feedback loop convergence based,
  - Develop destination choice models and eliminate *k*-factors, and
  - Further develop the freight model.

All the recommendations of the peer review have been addressed as of 2009, with the exception of the recommendation of a commercial vehicle survey.

An external cordon line video origin–destination study was conducted for TDOT and the TPO in 2007. The TPO and TDOT commissioned the 2008 East Tennessee Household Travel Survey covering an additional 1,400 households in Knox and Blount Counties as well as the collar counties of Loudon, Anderson, Union, Grainger, Jefferson, and Sevier to supplement the original survey in 2000 (*8*). Knoxville Area Transit also had a small onboard survey conducted in September 2008. All three additional data sources were incorporated in development of the new version of the KRTM.

Before the update of the KRTM and per the recommendation of the peer review, the TPO had an urban land use allocation model developed to assist in developing future land use scenarios. A short-term update of the KRTM completed in 2008 updated the model to a new 2006 base year and incorporated several of the more easily implemented improvements recommended by the peer review, such as making the feedback loop convergence based. Development of the advanced version of the KRTM addresses the remaining recommendations of the peer review with regard to model refinements, including development of mode and destination choice models and improvements to the freight model.

Development of an advanced travel model was partially motivated by the 2005 peer review but goes well beyond its recommendations. The move toward a new model design, incorporating experimental destination choice models was motivated by several factors.

The peer review commented on the *k*-factors in the previous version of the model and, more generally, it was recognized that gravity models were not performing well in the diverse, multinucleated Knoxville region. The hope that more sophisticated models, incorporating additional variables, could do a better job of replicating and predicting travel patterns in the region was a significant motive for the model upgrade. This hope seems at least partially vindicated by the results of the new model development. A comparison of the total daily trip tables produced by the old and new versions of the KRTM reveals that the new model provides a 33% increase in explanatory power over its predecessor (see the final section for details).

A seminar conducted for the TPO by its consultant team in 2007 and also attended by representatives of TDOT and the University of Tennessee helped identify additional planning issues of concern and provided further motivation for the new version of the model. In addition to the spatial distribution of trips, three broad issues received considerable attention. The first issue was the interaction of transportation with land use and the sensitivity of the travel model to different land use scenarios that might be developed with the help of the new land use model. There was some discussion of possible new land use development patterns or policies favoring more dense, mixed-use development. The possible exploration of transit-oriented development was also raised, and there was a general reaffirmation of the peer review's recommendation of developing mode choice models to support transit planning, although it was tempered by recognition of the limited transit mode share and resources for transit and transit planning. The possibility of future tolling or pricing scenarios also received significant attention, although developments in Tennessee may have now made those options less likely.

The TPO's whole data collection and model development program over the past decade supported development of the new advanced travel model. In particular, three elements were critical and are worth highlighting. First, collection of the household travel surveys provided the primary data source for development of the new model. The TPO could not have afforded to collect a 3,000-household survey in a single year. However, collecting two surveys in 2000 and 2008, which yielded a combined data set of roughly that size, proved feasible with the support of TDOT. Moreover, by collecting the data at two different times, the data set included not only the typical cross-sectional information but also a longitudinal dimension with observations at times when gas prices were very different. As a result it was possible to incorporate gas price as a variable in the new model; a single, traditional travel survey collected at one point in time would not have supported this inclusion.

Second, ongoing development of the trip-based model also supported development of an advanced model by providing a tested highway network, zone system, and external model. Although they may appear small, they represent a significant effort that, by being incorporated in earlier model development, made final development of the advanced model feasible.

Third, the model peer review in 2005 and the later visioning seminar in 2007 provided motivation and direction for development of the model and helped strengthen the collaboration between TDOT and the TPO. These events helped identify the accuracy of the spatial distribution, transit forecasting capability, and improved sensitivity to alternative land use scenarios as priorities for further model development. Many aspects of the new model's design developed in response to identification of the importance of these issues. To highlight one example, variables like the activity diversity within zones and intersection approach density were incorporated to provide greater sensitivity to land use and development patterns.

MPOs interested in moving toward advanced travel models would be well served to follow these three elements of the Knoxville TPO's program: programming ongoing data collection, making investments in maintaining and developing their current models, and periodically involving a wider group of peers and stakeholders.

OVERVIEW AND DESIGN OF KNOXVILLE'S ADVANCED MODEL

The KRTM predicts average weekday traffic volumes for all roadway classes of Knox and Blount Counties and major arterials and collectors in Anderson, Jefferson, Sevier, Loudon, Union, and Roane Counties and portions of Grainger County (Figure 1). The model's roadway network covers more than 6,600 lane miles over an area of

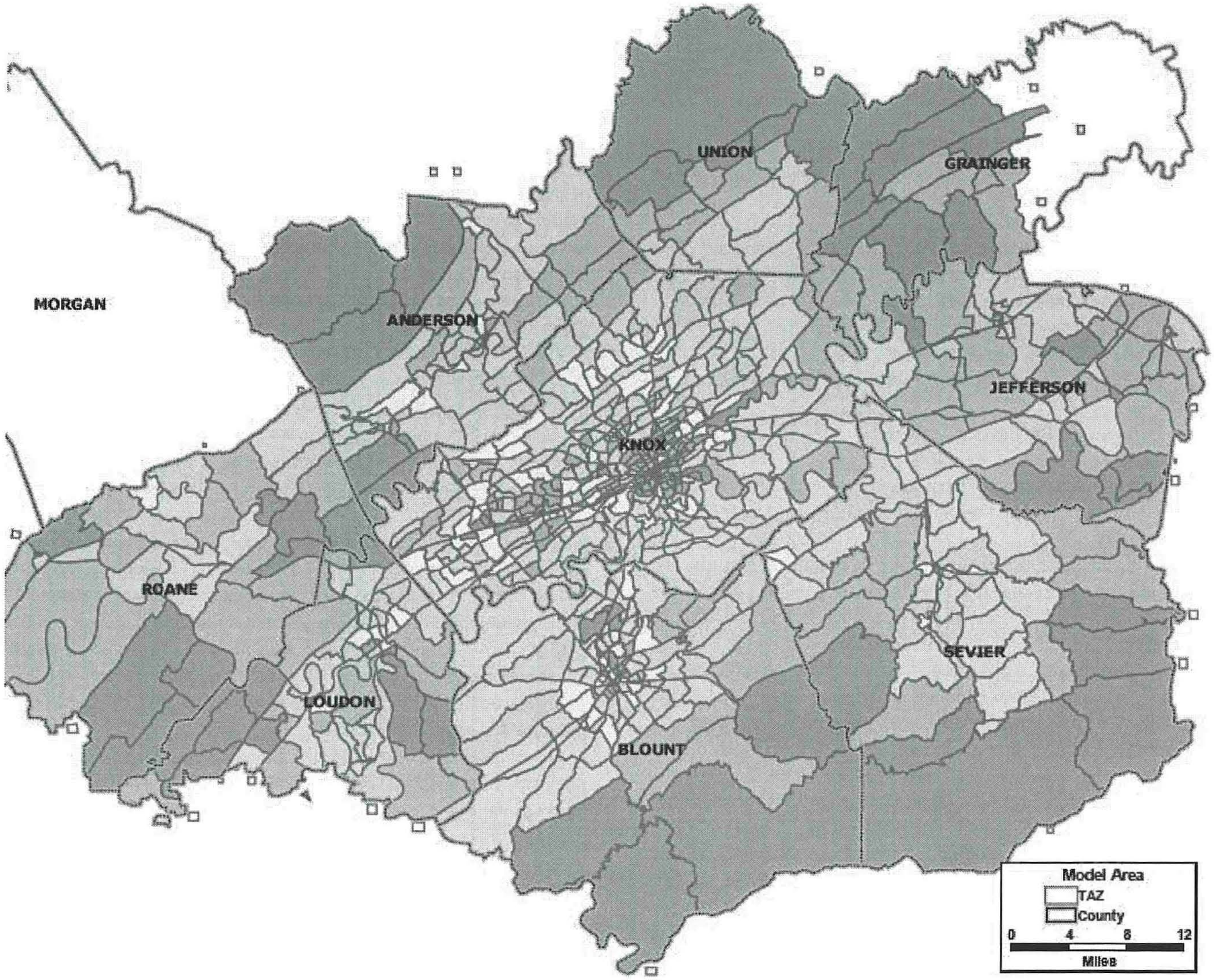


FIGURE 1 KRTM study area.

3,425 mi^2 represented by 1,019 traffic analysis zones (TAZs). The new KRTM also predicts the Knoxville Area Transit average weekday bus system ridership and number of average weekday bicycle and pedestrian trips within the region.

The prediction of roadway volumes, transit ridership, and bicycle and walk trips entails predicting the travel behavior of the region's population, which was estimated at 863,000 persons in 2006. The population is diverse and is distributed among several local activity centers, including Knoxville, Maryville, Oak Ridge, and the Gatlinburg–Pigeon Forge–Smoky Mountains tourism area. The area incorporates varied topography and includes a large student population associated primarily with the University of Tennessee.

The new KRTM represents the next generation of travel demand models. The previous version of the model was a good traditional model. The sequential trip-based design it implemented was based on research and practice that solidified in the 1970s and served as a standard for three decades. However, within the past decade there have been major advances in the ability to model urban and regional travel, which have been successfully used in about a half dozen metropolitan areas across the country. Both tour- and activity-based models as well as the hybrid trip–tour-based design implemented in Knoxville offer greatly improved policy sensitivity. In particular, the new KRTM offers the following features that its predecessor lacked:

- Sensitivity to fuel prices;
- Planning capability for transit, bicycle, and pedestrian modes;
- More realistic representation of special populations (seniors, low income, students);
- Sensitivity to urban design (mixed uses, grid versus cul-de-sac-style street networks);
- Ability to represent shifts in the timing of travel (e.g., due to congestion);
- Consistency with tours and trip-chaining behavior;
- Improved traffic impacts due to halo effects around major developments (e.g., malls);
- More accurate commuting patterns from destination choice models;

- Improved representation of speeds and delays from traffic signals, stop signs, and so forth;
- Improved accuracy of alternative analysis from new assignment algorithms; and
- Reduction of aggregation bias, which can skew model results.

The first tour- or activity-based models that offered these planning capabilities took considerable resources to develop and run. Most activity-based models took years to develop and run for 24 to 48 h on computers with 10 or more processors (*9*). In contrast, after several delays, primarily for the completion of data collection efforts, the new KRTM was developed in essentially 9 months and runs in less than 4 h on a standard dual-core laptop.

The speed of the new KRTM is a result of its hybrid model design. This architecture is based on recent research (*3–5*) and combines some elements of traditional "four-step" as well as several components from recent activity-based models but is ultimately distinct, made possible by the stop location and sequence choice structure original to the hybrid design.

In summary, the KRTM modeling process, illustrated in Figure 2, begins by generating a synthetic population of individual households based on the aggregate characteristics of the population encoded in the TAZs. Then a model predicting households' level of vehicle ownership is applied. The number of tours (sojourns beginning and ending at home) for various purposes (work, school, other) and the number of stops on those tours are predicted for each household. The dominant mode of travel (private automobile, school bus, public bus, walking, biking) is chosen for the household's tours for each purpose. Then, for automobile tours, grouping households within the same TAZ together in two basic market segments, probable locations of the stops on automobile tours are chosen. Next, for each probable stop location, a preceding location is chosen so that the resulting probable sequences of stops form tours that begin at home and proceed from one stop to the next until returning home. For each trip in the resulting travel pattern, the probability of walking, driving alone, or driving with passengers is predicted, as is the departure time (in 15-min periods) and toll eligibility. Finally, the trips are assigned to the roadway network and routes are chosen so that travelers minimize their travel time and costs. The resulting travel times are used to recalculate accessibility variables, and both are then fed back and used to repeat the process, beginning from the generation of tours and stops until the changes from one iteration to the next in the resulting roadway volumes are minimal. Many component models in this framework are familiar from traditional trip-based or standard activity-based models. The following sections provide additional details on some of the more innovative components.

The adjective "accessibility-based" originally used to describe the model design refers to the way accessibility variables are calculated and iteratively fed back through the modeling process to provide sensitivity to lower-level decisions in upper-level choices. Although different in some details, it is essentially the same as using logsum and proxy variables for the expected utility of subsequent choices in activity-based models. Earlier work (2) in development of the new

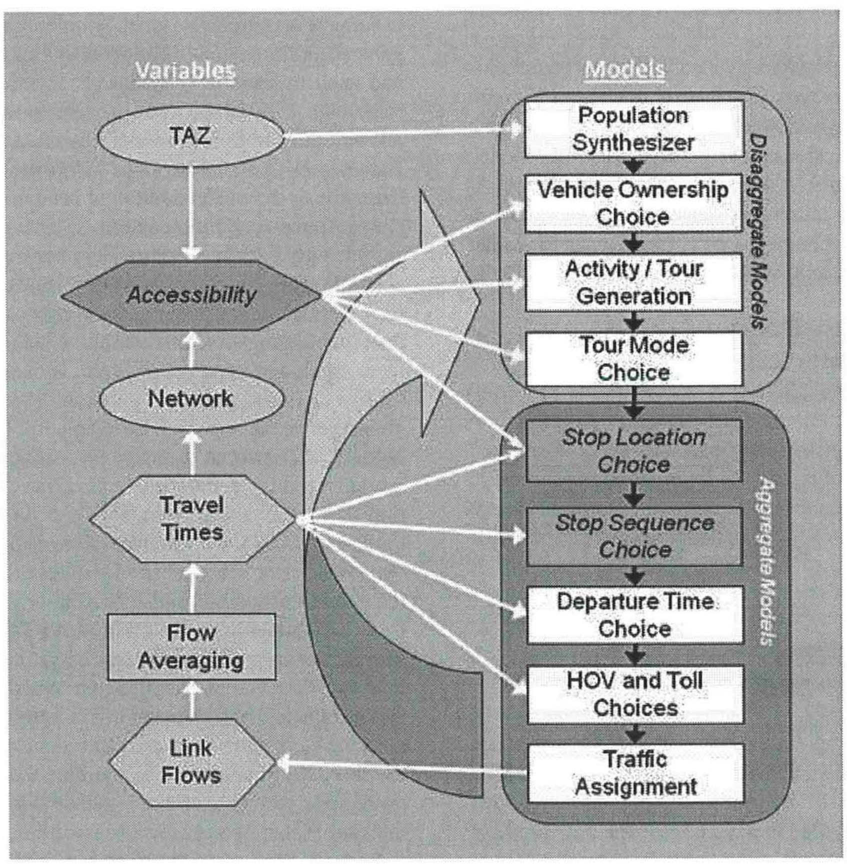


FIGURE 2 KRTM hybrid accessibility-based design.

model design characterized the architecture as accessibility based. However, in response to confusion between accessibility based and activity based and the helpfulness of the term "hybrid," the model design is now more frequently being identified by that term.

Hybrid refers to two ways the new model design blends aspects of four-step and activity-based models and defies traditional categorization. First, the hybrid KRTM model can be described as trip based insofar as it essentially produces aggregate trip table matrices of trips between origins and destinations rather than disaggregate records detailing individual travelers' activities. However, hybrid models like the new KRTM can also be described as tour based as the travel patterns they predict can be mathematically proven to be consistent with tours (*3*) and all travel is segmented within the model by types of tours; this process has the effect of eliminating non-home-based trips, which are problematic in traditional models. Hence, models of this design are hybrid trip-based and tour-based models.

Models like the KRTM are hybrid aggregate–disaggregate models. Unlike four-step models, which were (traditionally) entirely aggregate, and activity-based models, which are entirely disaggregate, the KRTM and similar models include both aggregate and disaggregate component models. Yet, despite its inclusion of disaggregate choice models, there are no random number draws or Monte Carlo simulation in the KRTM. As a result, the KRTM's model can produce a forecast of average or expected values from a single run, unlike activity-based or other simulation models, which require multiple runs to produce an average forecast. Any difference between two KRTM model runs is directly attributable to differences in their inputs as with traditional trip-based models. In simulation models, the multiple model runs are necessary when comparing alternatives to ensure that the difference between model runs results from differences in the alternative inputs rather than from differences in the random numbers drawn for each run.

The shift from the disaggregate framework of individual households to the aggregate framework of trips between zones midway through the model distinguishes the hybrid approach. The use of disaggregate components minimizes aggregation bias in the early steps of the model, including the particularly sensitive primary or tour mode choice. At the same time, the approach minimizes model run times by taking advantage of the fact that it is computationally much easier to predict a set of trips, which is consistent with tours, than to enumerate the individual tours.

The hybrid approach adopted here has limitations. It lacks the explicit representation offered by activity-based models of the interactions among household members and of constraints in timing travel and activities (although these phenomena are still implicit in this framework). However, given lower development costs and run time and the reproducibility of results, the hybrid model architecture presented a practical and cost-effective way to incorporate more sensitivity and realism in the KRTM to address the TPO's current and future planning issues.

INNOVATIVE COMPONENTS AND ADDED POLICY SENSITIVITY

This section presents some unique and innovative components and features of the KRTM's hybrid model design, highlighting sensitivity to policy variables. This sensitivity to different planning scenarios and policy alternatives provided motivation for the new model's development and is a key element of the model's added value over its traditional trip-based predecessor.

Population Synthesis and Vehicle Ownership

The KRTM's model process begins with generation of a synthetic population of households. Although various techniques have been used and are constantly being improved upon, this procedure has become fairly standard in activity-based models. The technique applied in the KRTM uses ordinal logit models to develop marginal distributions as inputs to iterative proportional fitting. The main difference between the KRTM procedure and those used in most activity-based models is that it does not use the resulting matrix as probabilities and randomly draw households from it to create a synthetic population; instead it uses the elements of the matrix as weights to produce a synthetic population deterministically. The vehicle ownership model that is then applied is a disaggregate, ordered-response logit model very similar to those used in activity-based models. The only difference is that it is not applied using Monte Carlo simulation but to produce an updated synthetic population segmented and reweighted by vehicle ownership.

Tour and Stop Generation

The new KRTM generates the number of tours and stops rather than full-day activity patterns or the number of home-based and non-home-based trips. The number of tours and stops of each type is estimated by using multiple regression models applied to the disaggregate synthetic population of households. First, the number of tours of each type is estimated for each household. Then, for each stop type, the number of stops per tour is estimated.

In this framework, the modeled behavior is dominated by the tour generation equations, with stop generation playing a secondary role, in some ways similar to, albeit simpler than, the day pattern activity generation framework developed by Ben-Akiva and Bowman (*10*) and used in most activity-based models, which allows for more trade-offs. This simplified approach to tour and stop generation is not essential to hybrid model design, and more elaborate model frameworks that allocate stops to tours may be developed at a later date, giving the model additional behavioral fidelity. However, the simple framework adopted for this version of the KRTM still offers significantly improved sensitivity over traditional models.

While cross-classification models were once viewed as an advance over regression models for generating trips, it was due to their ability to reduce aggregation bias compared with regression models, which were applied to zones as a whole. By applying regression models instead to a disaggregate population, aggregation bias is eliminated altogether in the approach adopted here. In this context, regression models offer two advantages over traditional cross-classification models used for generating trips. First, they allow the incorporation of additional variables. While cross-classification models are limited to two or three variables, regression models can include more variables, introducing sensitivity in resulting trip rates to gas prices, the number of seniors, and accessibility variables in addition to the basic demographic characteristics. As Table 1 shows, the tour and stop generation models offer sensitivity to considerably more variables than traditional cross-classification models. Each of these variables had a statistically significant effect and offers intuitive behavioral plausibility. Moreover, these effects enjoy the statistical support of the full data sample used to estimate the models and do not have problems common to cross-class methods, which result in some trip rates supported by very few observations.

Second, the use of regression models allows for limiting the nonlinearities in the model's travel rates to the two non-full-day

TABLE 1 Factors Affecting Tour and Stop Generation

	HH Workers	HH Non-workers	HH Students	Seniors in HH	HH Vehicles	HH Income	Gas Price	Accessibility
Work tours	+	N/A	N/A	–	N/A	N/A	N/A	+
Work stops	+	N/A	N/A	–	N/A	+	–	+
Other stops	+	–	+	–	N/A	+	–	+
School tours	N/A	+	+	N/A	N/A	N/A	N/A	N/A
School stops	N/A	N/A	+	N/A	N/A	N/A	N/A	N/A
Other stops	+	N/A	+	N/A	N/A	+	N/A	N/A
Other tours	N/A	+	N/A	N/A	+	+	N/A	+
Short maintenance stops	N/A	+	N/A	N/A	+	+	N/A	+
Long maintenance stops	N/A	+	–	+	+	+	N/A	+
Discretionary stops	N/A	+	N/A	N/A	+	+	–	+

NOTE: + variable (column) increases tour/stop rate (row); – variable (column) decreases tour/stop rate (row); HH, household; N/A, not applicable.

activity patterns and linear effects with plausible behavioral explanation: satiation effects (e.g., decreasing marginal increase in trips for each additional household member) and interaction effects (e.g., vehicles and workers increasing together increase travel more than either increasing by itself). Some satiation effects were incorporated in tour generation equations through the use of logarithmic transformations. Although interaction effects were widely tested, the only interaction effect that proved statistically significant was the interaction of gas prices and household income; increasing gas prices decreased certain stop rates but only for low-income households.

Including accessibility variables in tour and stop generation makes the trip rates in the KRTM elastic with respect to travel costs with modeled elasticities ranging from 0.13 to 0.16 for various tour and stop types. Introduction of this elasticity results in a variety of plausible effects, including less tour making by residents of rural (lower accessibility) areas, decreasing trip making in response to congestion (decreasing accessibility), induced trip making in response to added network capacity (increased accessibility), and induced stop making or halo effects in response to new land use developments in nearby zones (increased accessibility).

Tour Mode Choice

In the new KRTM, as in activity-based models, the mode of travel is modeled in two stages: tour mode choice and trip mode choice. First, after tours are generated, they are assigned a primary mode by tour mode choice models. Later, after the spatial distribution of stops creates trips, individual trips are assigned a mode, based on the primary mode of the tour, in trip mode choice models.

The KRTM predicts a household's choice of primary mode for its tours of each type using disaggregate nested logit models. The models incorporate level-of-service variables for each mode and fairly traditional cost and demographic variables including bus fares, gas prices, vehicle ownership, income, number of workers and students, and the presence of seniors. They also include built environment factors as determinants of walking. Several such variables proved significant for one or more tour types, including the percentage of sidewalk coverage, activity diversity (mixed-use development patterns), and intersection approach density, measuring the street network design (with high values for traditional grids and low values for cul-de-sac-style neighborhoods).

The incorporation of behaviorally sensitive tour mode choice models in the new KRTM represents significant added value compared with the previous model in which mode shares were fixed and totally insensitive to demographics, levels of service, cost, and other policy variables. In addition to automobile trips by occupancy class, the new model produces the system-level transit ridership, the number of transit trips generated by each residence zone, and the total regional number of daily walk and bike trips. Moreover, the model architecture allows for the straightforward addition of future component models to produce transit and nonmotorized trips at the route and street level.

The main innovation—and key difference between the tour mode choice models developed for the new KRTM and those common in activity-based models—is the way they measure the level of service provided by each competing mode and the related assumption of the hierarchy of travelers' choices (i.e., whether travelers' destination choices depend more on their mode choices or vice versa). In the KRTM, the level of service provided by each mode is measured by accessibility variables. These variables can be understood as proxy variables for the expected utility of a subsequent destination choice, implying destination or stop location choice conditional on tour mode choice. This situation represents the reverse assumption from that typical in traditional trip-based and standard activity-based models.

The traditional hierarchy, with mode choice conditional on destination choice, was first developed for very large metropolitan areas with significant choice rider markets, and it is likely appropriate for modeling affluent commuters in this context. However, this model structure implies that travelers are generally more likely to change mode than destination. This assumption may be incorrect for many travel markets in the United States and may partially account for the well-documented optimism bias in transit forecasts (*11*). More recent research abroad (*12, 13*) and evidence from the Knoxville data set (*5*) suggest that the KRTM reverse hierarchy and its opposite assumption that travelers are more likely to switch destinations than modes is more appropriate. In general, this assumption appears more plausible in markets like Knoxville with few choice riders where mode choice is generally a foregone conclusion (i.e., either travelers have access to a car and do not consider riding transit or they do not have access to a car and rely on transit, thus choosing their destinations, possibly even workplace, based on where the transit system can get them).

Stop Location and Sequence Choice

The KRTM hybrid model design hinges on its double destination choice framework of stop location and sequence choice models. It is the first model to apply this structure in practice, although it has

previously been the subject of research (*3, 4*). The double destination choice framework offers a substantial improvement over traditional trip-based models.

The spatial distribution of trips in traditional models, based on a single gravity model for each trip purpose, is open to several serious critiques. First, it is insensitive to trip-chaining efficiencies (e.g., the tendency of travelers to group their stops into convenient tours, such as stopping at restaurants near their workplace and frequent shopping locations). However, it has been demonstrated that this situation can be remedied with agglomerating and competing destination choice models, which incorporate variables measuring each destination's accessibility to locations with complementary activities (*4*). Trip-chaining efficiencies thereby incorporated as destinations that are more accessible to complements are more convenient and therefore more likely.

Most crucially, however, traditional trip distribution models can produce travel patterns that are physically impossible, because they are not consistent with the basic physical requirement that travel takes place in continuous paths through space and time and in the context of daily travel, closed tours. This is a serious problem with traditional models. The reason for it is surprisingly straightforward. Traditional models produce trips (trip tables) as a result of two models, trip generation and trip distribution, representing behaviorally the choice to participate in an out-of-home activity and the choice of where to engage in this activity (e.g., the choice to go shopping and the choice of where to shop). However, it will produce trips only if the home is assumed to be the trip's origin. In the general case including non-home-based trips, two spatial choice models (such as gravity and destination choice models) are necessary to define a trip by its origin and destination.

The double-destination choice framework adopted in the KRTM addresses the choice of both a trip's destination and its origin. First, in stop location choice, travelers choose all the destinations and locations where they will stop during the day—where they will go. Next, in stop-sequence choice, travelers choose an origin for each destination they will visit—where they will go from. The choice of origins must simply obey the constraint that each place that travelers visit is an origin exactly as many times as it is a destination. This "traveler conservation constraint" requires that as many travelers arrive at as leave each location every day so that travelers are never created or destroyed in the model. This constraint, together with the basic structure of the model, ensures that it will produce physically possible trips consistent with closed tours. For a formal demonstration of this process, see Bernardin (*3*).

The KRTM stop-location choice models address not only the incorporation of convenience and trip-chaining efficiencies but other effects as well, including the following:

- The effects of various impedances including
  - Travel times and
  - The psychological barrier represented by political boundaries and river crossings;
- The effects of traditional attraction or size variables such as employment and enrollment;
- The effects of other destination qualities, such as
  - Their accessibility to complements and to substitutes,
  - Their degree of activity diversity (mixed uses), and
  - The cost of parking; and
- Effects of travelers' characteristics, such as income and accessibility of their residence.

Most effects presented in Table 2 are incorporated in the model by adding terms to the logit model's utility function. However, the traveler heterogeneity effects related to income and residence location are handled differently. Analysis of average travel times from home to stop locations of the various types by income group indicated that

TABLE 2 Factors Affecting Stop Location Choice

	Impedance				Destination Qualities					
	Travel Time × Accessibility of Traveler's Residence	River Crossing	County Line Crossing	Intra-zonal	General Accessibility	Access to Complements	Access to Substitutes	Bus Availability	Activity Diversity	Pay Parking
Work Tours										
Work (low inc.)	–	–	–	+	N/A	N/A	N/A	+	N/A	N/A
Work	–	–	–	+	+	N/A	N/A	N/A	N/A	N/A
College	–	N/A	–	–	N/A	N/A	N/A	N/A	N/A	N/A
Nonwork	–	–	–	+	N/A	+	–	N/A	N/A	N/A
UT Tours										
UT campus	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Other	–	N/A	–	+	N/A	+	–	N/A	+	N/A
School Tours										
School	–	–	–	+	+	N/A	N/A	N/A	N/A	N/A
Other	–	–	N/A	+	N/A	+	N/A	N/A	+	N/A
Other Tours										
Short maint.	–	–	–	–	N/A	+	–	N/A	+	–
Long maint.	–	–	N/A	+	+	N/A	N/A	N/A	N/A	N/A
Discretionary	–	–	–	+	+	+	N/A	N/A	+	N/A

NOTE: UT = University of Tennessee; K–12 = kindergarten through 12th grade; inc. = income.

the only statistically significant difference was between low income (<$25,000 per year) and other work locations. Income was therefore used simply to segment the model and estimate separate work location choice models for low-income workers and other workers.

Incorporation of the accessibility of travelers' residence location reflects the fact that when people choose a residence location, they also effectively choose how far they are willing to travel. Travelers who live in dense, urban (high accessibility) areas are likely to have shorter trip lengths than travelers who live in remote, rural (low accessibility) areas. In the stop location choice models developed here, travelers' willingness to travel—and hence, trip lengths—varies as a function of the accessibility of their residence. In many gravity models, a gamma function is used as the friction factor function. In many destination choice models, an exponential function of travel time (t) is used as the friction factor function ($f_{hj} = e^{\beta_t t_{hj}}$) so that the term in the utility simplifies ($\beta_t t_{hj}$) and the willingness-to-travel parameter, β_t, can be easily estimated. However, in the models adopted here, travel time is interacted with the accessibility of the residence zone (a_{0h}) so that the friction factor term in the utility function becomes $\beta_t a_{0h} t_{hj}$ (the friction factor would be given, $f_{hj} = e^{\beta_t a_{0h} t_{hj}}$). The results of model estimation, documented in detail elsewhere (*14*), support the general hypothesis that rural residents are willing to visit locations farther from their homes than urban residents. In general, the willingness to travel of residents of the most urban (most accessible) areas was about 10% lower than the regional average, whereas the willingness to travel of residents of the most rural (least accessible) areas was about 200% higher or twice the regional average for most stop types. Together with the sensitivity-to-travel costs introduced through incorporation of accessibility in tour and stop generation, the KRTM begins to present a more realistic picture of travel in different area types with residents of rural areas making fewer but longer trips than urban residents; in traditional models, all travelers behave the same regardless of their residence location.

The KRTM is also believed by the authors to be original in statistically estimating parameters related to the reluctance of travelers to cross major rivers and county lines [although turnpikes were used as a psychological barrier by Chow et al. (*15*), and origin county–destination county k-factors have been common]. The amount of additional impedance represented by these psychological boundaries varied significantly for different stop types. For the choice of work and school locations, county lines represented 3.5 to 3.6 min of additional travel time (based on the value of travel time for travelers residing in average accessibility areas); for other stop locations, their effect was not particularly significant, generally representing less than a minute of additional travel time. The significance for school location choice, in particular, stands to reason, as most students attend public schools within school districts that respect county lines. The tendency of workers to work in the same county where they live is less obvious but not implausible. The impact of river crossings also varied in magnitude by stop type, generally representing between 0.9 and 3.3 min of additional travel time, with the higher values for school and work location choice but, unlike county line crossings, was almost universally significant.

The KRTM remaining component models for predicting trip mode choice, departure time choice, toll eligibility, and route choice or traffic assignment offer some unique improvements as well but are largely comparable to components of other travel models. The main features of note are the use of 15-min time periods and sensitivity to peak spreading in departure time choice.

BENEFITS AND COSTS OF NEW KRTM

The new hybrid KRTM offers all the improvements in behavioral fidelity and added sensitivity to policy variables detailed in the preceding sections at considerably lower cost than previous advanced

Destination Size (attractions)							
Basic Employment	Industrial Employment	Retail Employment	Service Employment	University Employment	UT Student Residents	K–12 Enrollment	HH or HH Population
+	+	+	+	N/A	N/A	N/A	N/A
+	+	+	+	N/A	N/A	N/A	N/A
N/A	N/A	N/A	N/A	+	N/A	N/A	N/A
N/A	N/A	+	+	N/A	N/A	+	N/A
N/A	N/A	N/A	N/A	+	N/A	N/A	N/A
N/A	N/A	+	+	N/A	+	N/A	N/A
N/A	N/A	N/A	N/A	N/A	N/A	+	N/A
N/A	N/A	+	+	N/A	N/A	+	+
N/A	N/A	+	+	N/A	N/A	N/A	N/A
N/A	N/A	+	+	N/A	N/A	N/A	N/A
N/A	N/A	+	+	N/A	N/A	N/A	+

models. As noted earlier and as others have reported (*9*), most activity-based models

- Typically take 2 to 3 years to develop,
- Cost $600,000 to $800,000 in consultant fees (not including data collection),
- Run for 1 or 2 days,
- Run on custom-developed software, and
- Run on computers with 10 or more processors.

TRB's *Special Report 288* stated, "Insufficient evidence exists that advanced models can be implemented for a reasonable cost and will provide significant improvements over current practice. . . . There are valid concerns about the costs associated with the new models and the amount of data needed to specify, calibrate, and validate them." (*1*, p. 6).

Compared with other advanced models, costs of the KRTM to the Knoxville TPO have been modest. The new KRTM

- Was developed in 9 months after several delays primarily related to data collection,
- Cost less than $300,000 in consultant fees (not including data collection),
- Runs in less than 4 h,
- Runs on commercially available software (with custom macros), and
- Runs on a standard laptop with a 2-GHz dual-core processor.

As noted, the KRTM hybrid approach makes some compromises and has limitations compared with activity-based approaches—most notably, the lack of explicit modeling of intrahousehold interaction and constraints on the timing of activities and travel. However, it offers substantial improvements over traditional trip-based models and at relatively low cost.

Attempts to compare the performance of activity-based and traditional trip-based models have often highlighted the need to consider differences that may bias the comparison. In comparing the previous trip-based and the new hybrid versions of the KRTM, several points are worth noting. The network and zone system used in the 2008 update of the trip-based KRTM and the 2009 hybrid KRTM are the same and both were validated to the same base year and set of counts. However, the trip-based KRTM was based on only the 2000 household survey data; the new hybrid KRTM was developed from the combined 2000 and 2008 household survey data. Both models used the same external–external trip table developed from the 2007 cordon line survey, but the new KRTM includes an improved external–internal trip model, a new visitor travel model, and improvements to the truck model in addition to improvements to the core component models presented here.

Keeping these factors in mind, some comparisons can be made between the previous, trip-based KRTM and the new hybrid KRTM. Experience with other advanced models has shown that base year assignment validation statistics are not necessarily the best means of comparison as traditional trip-based models can be calibrated to reproduce an observed base year quite well, which was arguably the case with the Knoxville models as well, although the new advanced model did result in a somewhat better assignment of 28.13% root-mean-squared error (RMSE) versus 32.95% RMSE for the traditional model. Assignment models can cover a multitude of errors in the underlying trip tables, however, so a more meaningful and revealing comparison is between the trip tables produced by the models. Comparing the total daily person trip tables from the two models with the observed person trips from the household surveys, the McFadden rho squared (versus the null hypothesis of total randomness) (*16*) for the new model was 0.168 for the new model versus 0.126 for the old model. Thus, although a great deal of the observed spatial variation in trips remains to be explained, the new hybrid model offered a 33% increase over the explanatory power of its predecessor.

Sensitivity analyses and comparisons of model predictions with observations of conditions other than the base year data used to calibrate the models are necessary to compare performance of the old and new KRTM as planning and forecasting tools. However, the sensitivity of the new model to variables such as gas prices, accessibilities, parking costs, the senior population, activity diversity, intersection approach density, and other variables to which the old model was insensitive gives some basis for expecting better forecasts from the new model.

REFERENCES

1. *Special Report 288: Metropolitan Travel Forecasting, Current Practice and Future Direction.* Transportation Research Board of the National Academies, Washington, D.C., 2007.
2. Bernardin, V. An Accessibility-Based Approach to Travel Demand Forecasting: New Alternative to Four-Step and Activity-Based Methods. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C., 2008.
3. Bernardin, V., Jr. *A Trip-Based Travel Demand Framework Consistent with Tours and Stop Interaction.* Dissertation. Northwestern University, Evanston, Ill., 2008.
4. Bernardin, V., F. Koppelman, and D. Boyce. Enhanced Destination Choice Models Incorporating Agglomeration Related to Trip Chaining While Controlling for Spatial Competition. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2132,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 143–151.
5. Newman, J., and V. Bernardin. Hierarchical Ordering of Nests in a Joint Mode and Destination Choice Model. *Transportation,* Vol. 37, No. 4, 2010, pp. 677–688.
6. NuStats. *2000 Household Travel Behavior Study, Final Report.* Knoxville Regional Transportation Planning Organization, Knoxville, Tenn., 2001.
7. Everett, J. *Knoxville Travel Demand Model Peer Exchange.* The University of Tennessee Center for Transportation Research, Knoxville, 2005.
8. NuStats. *2008 East Tennessee Household Travel Survey, Final Report.* Knoxville Regional Transportation Planning Organization, Knoxville, Tenn., 2008.
9. Rossi, T. F., B. Winkler, T. Ryan, K. Faussett, Y. Li, D. Wittl, and M. A. Zeid. Deciding on Moving to Activity-Based Models (or Not). Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
10. Ben-Akiva, M. E., and J. L. Bowman. Activity-Based Disaggregate Travel Demand Model System with Daily Activity Schedules. Presented at EIRASS Conference on Activity-Based Approaches: Activity Scheduling and the Analysis of Activity Patterns, Eindhoven, Netherlands, 1995.
11. Flyvbjerg, B., N. Bruzelius, and W. Rothengatter. *Megaprojects and Risk: An Anatomy of Ambition.* Cambridge University Press, Cambridge, United Kingdom, 2003.
12. Abrahamsson, T., and L. Lundqvist. Formulation and Estimation of Combined Network Equilibrium Models with Applications to Stockholm. *Transportation Science,* 1999, Vol. 33, No. 1, pp. 80–100.
13. Debrezion, G., E. Pels, and P. Rietveld. Modelling the Joint Access Mode and Railway Station Choice. *Transportation Research E,* Vol. 45, No. 1, 2009, pp. 270–283.
14. Bernardin, Lochmueller & Associates, Inc. *2008–2009 Knoxville Regional Travel Model Update, Final Report.* Knoxville Regional Transportation Planning Organization, Knoxville, Tenn., 2009.
15. Chow, L.-F., F. Zhao, M.-T. Li, and S.-C. Li. Development and Evaluation of Aggregate Destination Choice Models for Trip Distribution in Florida. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1931,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 18–27.
16. Windmeijer, F. A. G. Goodness-of-Fit Measures in Binary Choice Models. *Economic Reviews,* Vol. 14, 1995, pp. 101–116.

The Transportation Planning Applications Committee peer-reviewed this paper.

Estimating Price Elasticities of Ferry Demand

Thomas Adler, Youssef Dehghani, and Celine Gihring

Washington State Ferries operates what is by far the largest ferry system in the United States. The ferry routes serve as the primary transportation link to several islands in Puget Sound. Customers include a mixture of commuters, nonwork travelers, and tourists who use the ferries as walk-on or drive-on passengers. The system experiences seasonal and time-of-day peaks that result in some routes operating at capacity for drive-on customers. As part of a recently completed long-range planning effort, Washington State Ferries in association with the state's transportation commission wanted to evaluate the impact of changes in fare policies—in particular, to analyze the effects of charging different fares at different times of day. Fare levels have changed several times over the past several years. The resulting elasticities corresponding to these fare increases were calculated and provide a reasonable foundation for estimating the effects of future fare changes. However, the elasticities do not reflect the effects of charging different fares by time of day. To estimate those elasticities, a stated preference survey was designed and administered to current drive-on customers. Data from the survey were used to estimate both segment-level discrete choice models for choice of alternative drive-on sailings, walk-on, or shifting to an alternative route or mode. Hierarchical Bayes estimation was used to develop models that reflect random heterogeneity in preferences and those models were used in a simulation model to estimate time-of-day fare elasticities. This paper describes the resulting fare elasticities and their implications for fare policy.

The Washington State Department of Transportation Ferries System (WSF) was formed in 1951 and has grown to 10 routes with more than 500 daily sailings. Its vessels carry nearly 23 million riders annually, making the system by far the largest in the United States. Figure 1 shows the routes on the system, most of which are served by vessels that carry both automobile (drive-on) and passenger (walk-on) customers.

Because of continuing growth in the Puget Sound region, the demand for ferry service is projected to increase as the population in ferry-served markets grows. However, the system is constrained by tight financial resources, limited vehicle-carrying capacities to accommodate peak demand, and aging vessels and terminals. The "ferry bill" passed by the legislature specified that a long-range planning effort should be undertaken, supported by ferry travel survey data and improved technical approaches in ridership to address these issues. Ridership modeling–forecasting was a key analytic component to support development of the new long-range plan. It provided key input in a variety of areas, including assessing future demand for ferry services, shifting demand between travel modes, testing pricing strategies, and accommodating growth in demand while shifting riders into time periods that have excess capacity.

The final plan, completed in summer 2009, provides guidance to WSF future service and investment decisions through the year 2030. In particular, the legislation included the following directives for WSF:

- Reconnect with customers to get better information about their travel,
- Improve its forecasting approach to ensure its plans are based on the best projections of future needs,
- Develop strategies to minimize costs,
- Implement adaptive management practices to keep costs as low as possible while continuously improving the quality and timeliness of services,
- Consider operational and pricing strategies that would improve asset utilization and reduce costs, and
- Reestablish the vehicle level-of-service standard to fit better with current policy and funding realities.

The ferry bill posed a challenge as well as an opportunity in the area of ridership forecasting analysis. The joint technical committee's ferry finance study has raised concerns related to the methods previously used to develop ridership forecasts.

A technical working group was formed to assist in developing and reviewing the forecasting methods. The group included expertise from public agencies and consulting firms in the area of travel demand modeling and forecasting analysis. This group met regularly over the course of about 1.5 years. They reviewed the ridership forecasting approaches, underlying data and assumptions, and model results and provided valuable feedback to the modeling team in the form of recommended refinements to procedures and "reasonableness" assessments of the results.

The resulting plan considers a variety of possible strategies related to pricing. Specifically, the ferry bill required that "the department shall annually review fares and pricing policies applicable to the operation of the WSF." The WSF is required to develop fare and pricing policy proposals that must accomplish the following:

- Recognize the uniqueness of travel sheds in terms of fare-box recovery rates and pricing policies;
- Use data from the market survey conducted by the Washington State Transportation Commission;

T. Adler, Resource Systems Group, Inc., 55 Railroad Row, White River Junction, VT 05001. Y. Dehghani, Parsons Brinckerhoff, Inc., 999 Third Avenue, Suite 2200, Seattle, WA 98104. C. Gihring, Washington State Department of Transportation Ferries Division, 2008 East Crescent Drive, Seattle, WA 98112. Corresponding author: T. Adler, tadler@rsginc.com.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 59–66.
DOI: 10.3141/2176-07

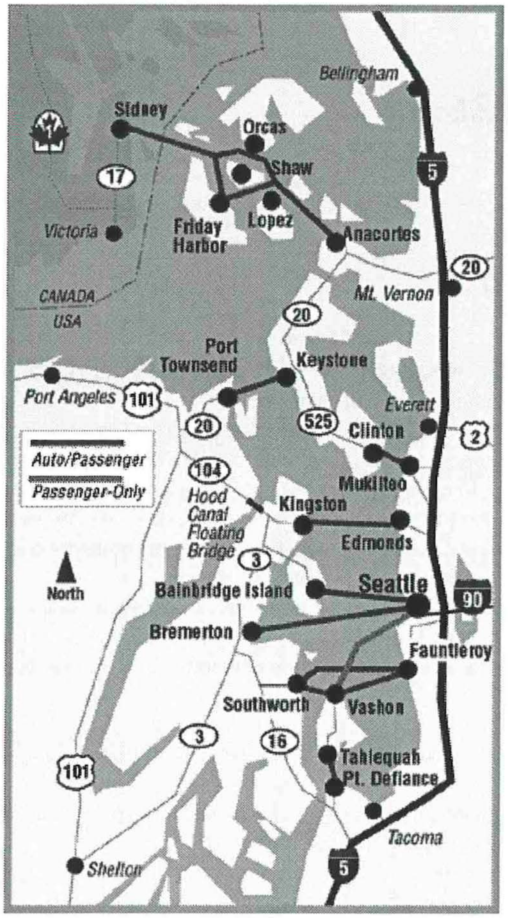


FIGURE 1 WSF routes.

• Be based on input from public hearings with affected ferry users and review with affected ferry advisory committees;
• Generate the amount of revenue required by the biennial transportation budget;
• Consider the impacts on users, capacity, and local communities; and
• Keep the fare schedules as simple as possible.

Also, the WSF was asked to consider options for using pricing to level vehicle peak demand and to increase off-peak ridership. Key pricing strategies evaluated in the plan include the following:

• Peak period surcharges (i.e., applying a surcharge to a peak time of day, day of week, or season of the year),
• Off-peak discounts (i.e., using a pricing incentive to encourage travelers to use lower-demand sailings), and
• Passenger discounts (i.e., using a pricing incentive to encourage a shift from drive-ons to walk-ons).

Baseline ridership forecasts were used to determine where level-of-service standards were not being maintained. Evaluation of the potential effects of price strategies on revenue and ridership required using ridership forecasts, information on seasonal fluctuations in ridership, and fare elasticities. Baseline ridership forecasts were used in the plan to test effects of pricing strategies to manage demand and to shift demand from peak to off-peak periods and from vehicles to walk-ons.

This effort resulted in an updated and improved ferry forecasting analysis procedure that included estimation of fare elasticities, based on historic data and ferry customer survey data, to support the evaluation of pricing strategies. Historic data were available to estimate overall fare elasticities. As shown in Figure 2, average fare levels have varied over the 45+ years of operation. These variations are sufficient to allow estimates of overall fare elasticities.

However, these data do not describe the effects of different pricing structures, such as time-of-day pricing and differential pricing by type of customer, which were among the fare policies to be evaluated in the current plan. Using the results of an onboard customer survey conducted by the Washington State Transportation Commission, the consultant team estimated both overall fare elasticities and elasticities describing the effects of different fare policies. These elasticities were used in the plan to test potential effects of pricing strategies on shifting demand among modes and time periods. This paper describes the work that was completed to estimate these fare elasticities for current WSF passengers.

OBJECTIVES

The technical objective of the work described here was to estimate the fare elasticities associated with different types of changes in the structure of ferry rates. While fare elasticities have been estimated for

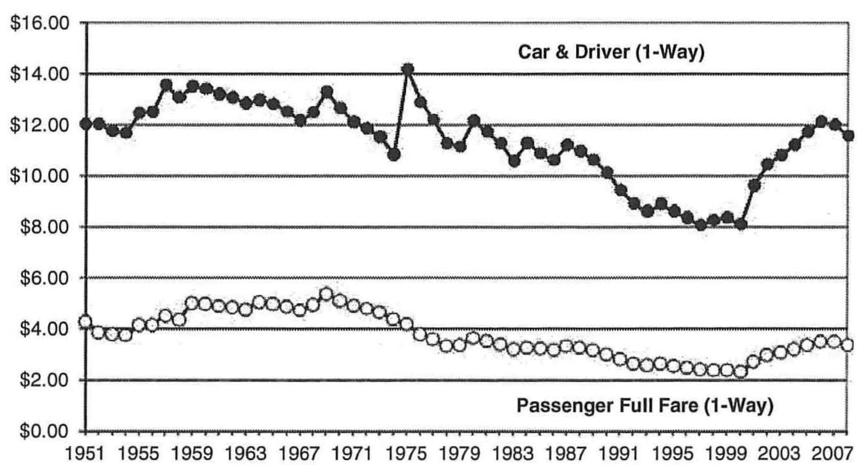


FIGURE 2 Average historic fare levels (2008 U.S. dollars; *1*).

a wide variety of public transit systems and of fare structures (*2, 3*), the WSF application is unique in several respects:

1. For many WSF passengers, the ferry is the only reasonable mode available for their trip.
2. Drive-on fares are quite high relative to the more typical transit commuting modes.
3. The fare structures to be evaluated include both time-of-day differentials and changes in the relative rates by type of customer (drive-on versus walk-on).

The policy objective supported by this work was to identify fare structures that could alleviate peak capacity constraints and assist in meeting the system's financial needs.

APPROACH

Because of the inherent limitations of the available historic data, a stated preference survey was designed to elicit information about WSF customers' likely response to fare and service changes. Stated preference methods have been widely used to estimate travelers' responses to new or changed transportation services (*4*). They are particularly useful if the changes to be evaluated are outside the range of past experience and for which historic data are of limited use. While stated preferences may not be perfect predictors of actual behavior, appropriate design of the survey instrument and careful analysis of the resulting data can produce estimates that closely approximate actual behavior.

For the WSF application, a set of stated preference exercises was designed for inclusion in a survey of predominantly peak-period drive-on customers. The survey was developed as part of a survey program administered by Opinion Research Northwest (ORC-NW). That program included a large-scale onboard survey from which respondents were selected for this online stated preference survey. The stated preference survey questionnaire included questions to determine details of the respondents' most frequent ferry trips for nondiscretionary (e.g., work) and discretionary (e.g., social, recreation) purposes. These trips were then used to set the context for stated preference scenarios in which fares and other service conditions are varied, creating realistic alternatives for making that trip. Respondents are asked to select the alternative they would most likely choose under those circumstances.

The use of a specific past trip as a point of reference is important in these surveys because travel decisions are commonly quite context specific; travelers have specific needs and constraints that vary considerably from day to day and from trip to trip and an average or typical trip does not reflect those real needs and constraints. By sampling across all trips made by respondents, a representative mix of these needs and constraints will be represented in the sample.

The WSF customers were asked to choose from five alternatives:

1. Drive-on the sailing chosen for the most recent trip,
2. Drive-on an earlier sailing,
3. Drive-on a later sailing,
4. Walk-on the sailing chosen for the most recent trip, and
5. Make the trip some other way or not at all.

Each alternative was described by a fare, a waiting time, and an actual sailing departure time. Fares, waiting times, and departure times for the earlier and later sailings were varied across scenarios. An example screen from the ORC-NW online questionnaire is shown in Figure 3.

Eight such scenarios were shown to each respondent, with the variables changing according to experimental design. The data from

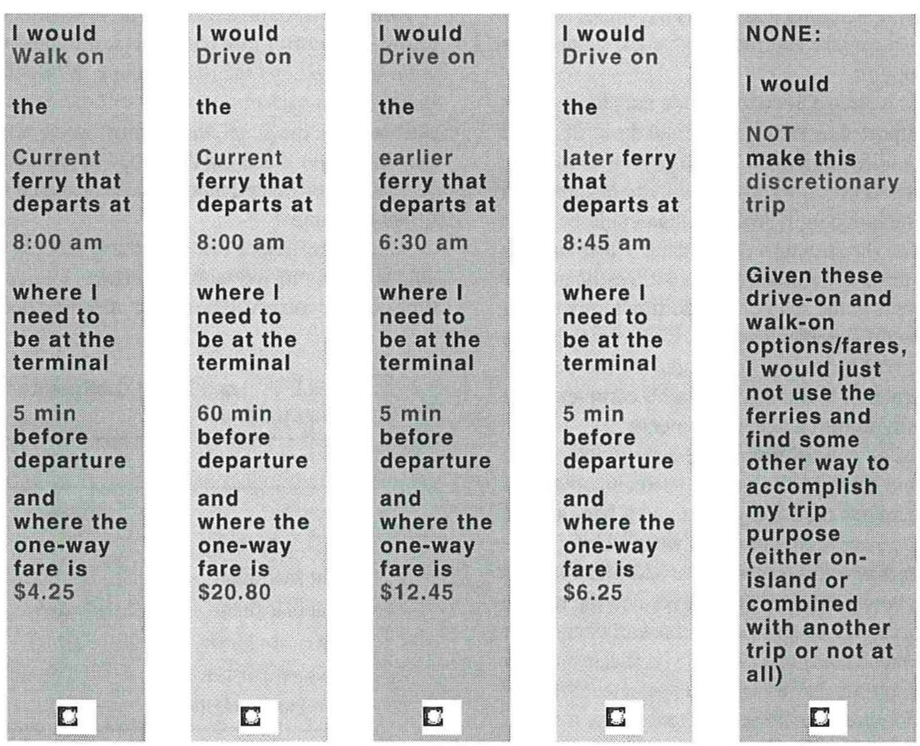


FIGURE 3 Example stated preference scenario.

the 838 respondents who completed the survey provide sufficient information to allow estimation of systemwide elasticities for current drive-on customers and in particular to determine how those customers would likely change their use of the system in response to changes in fare levels and policies.

MODEL DEVELOPMENT

Stated preference data are designed to support statistical estimation of travel choice models, which in this application are models to determine the effects of price and the other service characteristics on the likelihood of current WSF drive-on customers to choose each of the five travel alternatives detailed above. The basic structure used for these travel choice models is a multinomial logit form that has been widely used for these types of applications (*5*). The stated preference data were merged with demographic and trip data from the associated respondents. Together, these two sets of data allowed estimation of models that incorporated systematic and random heterogeneity.

It has been shown that some differences in preferences are systematic effects of demographic variables, such as income, and it is important to incorporate those effects in the model structure before modeling random differences in preferences (*6*).

The models were specified to include all the variables that were varied among the choice-based conjoint experiments in addition to selected demographic and trip characteristics. Fare and waiting and departure time differences (for the earlier and later sailings) were specified as nominal variables in dollars and minutes, respectively, and were treated as continuous rather than categoric variables.

Initial work conducted previously by others treated fare and time shift as categoric variables with seven and four levels, respectively. However, the resulting individual-level models displayed significant nonmonotonicity; well over three-quarters of the individuals in the sample had estimated values that suggested higher prices were preferred to lower prices. This finding was simply a result of stretching the data to estimate more fare and time shift coefficients than they could reasonably support.

In addition, the models were specified to include only the number of terms that could be supported by the experimental design and that followed econometric identification rules. Initial models developed previously by others were significantly overspecified, meaning that more coefficients were included in the models than could be independently estimated given the structure of the data. Those models had about 60 individual terms (coefficients) that were estimated for each of the 838 respondents in the sample. In addition to reviewing the data for outliers, several specification tests were conducted to determine the type of travel behavior represented by the data. In general, these tests indicated that the travelers who completed the survey responded carefully to the choice experiments.

The general form of the specifications has the three stated preference variables—fare, time between the sailing used and an earlier sailing time, and time between the sailing used and a later sailing time—and waiting time. In addition, constants were added for the walk-on and drive-on alternatives to represent the additional factors that might affect choice between these alternatives and the no-trip option. Separate constants were not specified for the earlier and later sailing options because those constants did not significantly differ from the one for the sailing time that was actually selected; the time shift variable adequately represents those differences. As a result, only two constants were specified for the five alternatives. At most, $n - 1$ constants can be independently identified for n alternatives because choices depend only on the difference in utilities between alternatives.

Fare and the time variables were treated as continuous variables and a number of linear and nonlinear functions were tested to determine the relationship between these variables and the utility—the economic term used to describe the general attractiveness of an alternative. In market research, the terms worth and partworth (for the contribution of a particular attribute) are more commonly used to describe the same quantity of the ferry alternatives. These tests indicated that the effects on utility do not significantly differ from linear within the range of fare values tested and similar tests indicated linearity in the effects of time shifts to earlier or later sailings. As a result, general linear specifications were used for these variables in all later work. The overall effect of utility in the multinomial logit model used for this work is nonlinear even with a linear representation within the utility function.

Systematic effects of demographic and other variables were explored through specification tests. The most consistent effect was found to be an income effect on price sensitivity. This effect has been noted in many other travel choice models and was incorporated here in a flexible nonlinear-in-the-parameters form. An additional exponent representing the elasticity of price sensitivity to income was estimated by a nonlinear search method (*7*), as opposed to a simple multiplicative interaction form commonly used in these models. As expected, systematic differences were found in the responses between discretionary and nondiscretionary trips. The sample was also segmented along three other dimensions to identify other systematic variations in preference. The following dimensions were tested:

1. Payment type: multiride discounted fare or full fare;
2. Actual sailing time: peak period or off-peak period; and
3. Route group: north, central, south, or island.

Although some differences in price sensitivity were found between multiride and full-fare passengers, the most significant differences were found among the route groups. While the sample sizes for the individual line groups were not sufficient to support detailed model estimation at that level, this segmentation structure was carried into the simulation modeling described in the next section.

The final estimated models and associated statistics are presented in Tables 1 and 2.

All the coefficient values in these models are intuitively reasonable when compared with other travel choice models. The fare and time coefficients are all negative, meaning that utility values decline

TABLE 1 Logit Model Coefficients for Discretionary Trips

	Coefficient	*t*-Statistic
Fare ($)	−0.136	−25.4
Shift earlier (min)	−0.0101	−19.9
Shift later (min)	−0.00962	−19.0
Wait time (min)	−0.0205	−14.5
Drive-on constant	3.04	33.5
Walk-on constant	0.679	11.5
Fare-income elasticity	−0.166	−5.0

NOTE: Observations = 4,170; initial log likelihood = −6,711; final log likelihood = −5,954.

TABLE 2 Logit Model Coefficients for Nondiscretionary Trips

	Coefficient	*t*-Statistic
Fare ($)	–0.126	–18.7
Shift earlier (min)	–0.0139	–19.6
Shift later (min)	–0.0136	–19.5
Wait time (min)	–0.0184	–12.0
Drive-on constant	2.9	25.8
Walk-on constant	0.87	13.1
Fare-income elasticity	–0.0918	–1.8

NOTE: Observations = 2,526; initial log likelihood = –4,065; final log likelihood = –3,606.

with increasing values and all are highly statistically significant. The income elasticities are negative, meaning that price sensitivity declines with higher incomes, as would be expected. Also as would be expected, these models indicate that travelers on nondiscretionary trips are less willing to shift to earlier or later sailings and are less cost sensitive than those making discretionary trips.

These results were compared with the model currently being used in the WSF forecasting system. The coefficient of fare is common between that model and the ones described in Tables 1 and 2. Travel time was not varied in this stated preference survey and so a comparable value is not estimated here. The fare coefficients estimated here are remarkably similar to that developed in 1999 for the current model: –0.129 in that model (after adjustments that are needed because of different model specifications) and –0.136 for discretionary and –0.126 for nondiscretionary trips.

These models describe the systematic differences that are most important in affecting choice among the ferry options, but there are also random differences among individuals that may be important. There are several ways to represent these random differences, the most common of which are use of mixed logit and hierarchical Bayes (HB) estimation (*8*). To maintain consistency with previous work that was conducted with these data, HB estimation was conducted using the model specifications described above. This process results in model coefficients for each of the individuals in the sample, representing their unique preferences. The results across the sample were similar to those shown in Tables 1 and 2 but with a different scale for the coefficient values. Because the scale of the resulting HB coefficients depends on the type of normalization used and other controls in the estimation process, the average scale was postnormalized to the aggregate multinomial logit coefficient scales.

ELASTICITY ESTIMATES

The price elasticity of demand is defined in economics as the ratio of the percent change in demand to the percent change in price. It is a measure of the relative responsiveness of demand to changes in price. Price elasticities are generally negative, meaning that as price for a service increases, demand for that service decreases. Services with price elasticities whose absolute value is greater than one are termed "elastic" and any increases in price for those services will result in decreases in both demand and gross revenue. Services whose price elasticities are less than one are called "inelastic" and increases in price for those services will result in reduced demand but higher gross revenue. The elasticities that result from the multinomial logit model used here are not constant but increase in absolute value with increasing prices. Thus, at some price, a service with inelastic demand will switch to elastic demand, implying that there exists a maximum gross revenue price point.

TABLE 3 Calculated Elasticities for Discretionary Trips

	Elasticity of Drive-On Volume to Drive-On Fares (10% fare increase)	Elasticity of Peak Drive-On Volume to Off-Peak Fares (20% off-peak fare decrease)
North routes	–0.40	0.74
Central routes	–0.31	0.65
South routes	–0.26	0.49
Island routes	–0.20	0.91
Overall	–0.30	0.64

The elasticity of ferry demand can be calculated by simulating the mode choice behavior of the population (as represented by the survey sample) with different fare structures. The individual-level HB models as described above were used in a spreadsheet simulation model to estimate elasticities under different fare policies. The individuals were weighted to be representative of the overall ferry population using weights calculated and supplied by ORC-NW.

Two initial pricing tests were conducted:

1. Ferry drive-on fares were increased by 10% and
2. Drive-on fares for the later and earlier sailings were decreased by 20%.

The arc price elasticities (those that in effect average the elasticities at the two points on the demand curve represented by the two price points) were calculated for the drive-on ferry population as whole and separately for each of the ferry route groups. The resulting arc elasticities are shown in Tables 3 and 4.

On the basis of these calculations, drive-on ferry demand is reasonably inelastic at current fare levels (first numerical column in Tables 3 and 4), which means that drive-on fares could be increased by at least small amounts and, while the resulting demand would decrease somewhat, gross revenues would continue to increase. Elasticities of discretionary trips are generally higher, and in particular they are more likely to shift to walk-on in response to higher drive-on prices. The resulting elasticities as estimated from the model of –0.30 for discretionary trips and –0.34 for nondiscretionary

TABLE 4 Calculated Elasticities for Nondiscretionary Trips

	Elasticity of Drive-On Volume to Drive-On Fares (10% fare increase)	Elasticity of Peak Drive-On Volume to Off-Peak Fares (20% off-peak fare decrease)
North routes	–0.43	0.59
Central routes	–0.37	0.52
South routes	–0.22	0.34
Island routes	–0.39	0.97
Overall	–0.34	0.51

trips bracket the actual observed historic price elasticity of −0.32, indicating that the stated preference-based models replicate actual behavior fairly closely.

The elasticities of peak drive-on sailings to reductions in off-peak fares (second numerical column in Tables 3 and 4) (they are technically cross-elasticities—the change in demand for one service as a result of a change in price of another service—which is why the signs are positive) are somewhat greater in absolute value, though still reflecting overall inelastic conditions. Here, as expected, the elasticity of nondiscretionary trips is somewhat lower than for discretionary trips, likely because there is less time-of-day flexibility in these trips.

In general, these results suggest the demand for drive-on boardings is somewhat sensitive to general increases in drive-on fares; a 10% increase would result in a more than 3% decline in drive-on boardings. However, the demand for peak drive-on sailings is even more sensitive to changes in fares for off-peak drive-on fares; a 10% decrease in off-peak sailing fares would result in a decline in peak sailing drive-ons of between 5% and 6%. Therefore, a differential time-of-day fare policy could result in significant reductions in peak drive-on demand levels.

APPLICATIONS OF ELASTICITIES

A spreadsheet-based sample enumeration model was developed to allow testing of a full range of pricing scenarios. The spreadsheet model allows fares to be specified by mode, by traveler segment, by time-of-day, and by route groups. It also is set up to facilitate comparisons among alternative scenarios in current or future years.

The spreadsheet model was used to calculate elasticities for a wide range of possible pricing scenarios to support development of the long-range plan. The spreadsheet model was used to determine the effects of fare changes on ferry demand. In particular, it was used to estimate the amount of change in passenger volumes that would result from different fare policy changes. Several types of fare policy changes were evaluated as part of the long-range planning effort:

- Peak-period fare surcharges in which fares for daily peak sailings were increased but all other fares remained the same. Table 5 shows the results from that analysis—the percentage increase is shown in the leftmost column and each row represents the estimated percent change in volumes for the given type of passenger (discretionary versus nondiscretionary trips) and for each trip type (walk-on, peak-period sailings, sailings before the peak, sailings after the peak, and passengers who choose not to make the ferry trip).
- Off-peak-period fare discounts in which the fares for daily off-peak sailings were decreased but all other fares remained the same. Table 6 shows the results from that analysis.
- Walk-on fare discounts in which the fares for walk-on passengers were decreased but all other fares remained the same. Table 7 shows the results from that analysis.
- Blended fare increases in which all fares were increased but the fares for walk-on passengers were increased at half the rate as those for drive-on customers. Table 8 shows the results from that analysis.

PRICING-RELATED RECOMMENDATIONS IN WSF LONG-RANGE PLAN

The final version of the WSF long-range plan contains detailed discussions of the pricing scenarios that were tested in this analysis. The time-of-day pricing alternatives showed significant promise as options for relieving peak-period capacity issues. However, based on input from public meetings and other considerations, the plan recommends, instead, use of a reservation system to manage demand for the limited peak-period drive-on capacity. The plan also discusses other pricing strategies, including time-of-day-based congestion pricing for "possible future consideration" after first implementing the reservation system. The plan notes that

> the pricing strategy with the greatest potential to shift travel behavior is congestion pricing. If reservations alone are not sufficient to shift demand then it may be necessary to evaluate a reservations plus variable congestion pricing approach. (*1*)

Similarly, the plan acknowledges that increasing the fare differential between walk-on and drive-on fares would reduce the demand for drive-on boardings and hence relieve capacity issues. However, the legislature specified that the differential not be increased. So, while these pricing options will likely not be exercised in the near term, they remain options for the future should additional measures be needed to manage ferry level of service more effectively.

CONCLUSIONS

The work described here was designed to estimate the effects of possible ferry fare policy changes. The stated preference method that was used to estimate fare elasticities appeared to produce reasonably

TABLE 5 Effects of Peak Ferry Fare Surcharges

	Discretionary Trips					Nondiscretionary Trips				
		Drive-On Board					Drive-On Board			
Peak Fare Surcharge (%)	Walk-On Board (%)	Peak (%)	Before Peak (%)	After Peak (%)	None[a] (%)	Walk-On Board (%)	Peak (%)	Before Peak (%)	After Peak (%)	None[a] (%)
10	3	−10	5	5	5	3	−9	5	5	5
20	5	−21	10	11	10	6	−19	10	10	9
30	8	−34	16	16	15	9	−30	15	15	14
40	10	−47	20	21	19	12	−42	20	19	18
50	13	−61	25	26	24	15	−55	24	24	22

[a]Includes those ferry riders who would make the trip some other way or not at all.

TABLE 6 Effects of Off-Peak Ferry Fare Discounts

	Discretionary Trips					Nondiscretionary Trips				
		Drive-On Board					Drive-On Board			
Off-Peak Fare Discounts (%)	Walk-On Board (%)	Peak (%)	Before Peak (%)	After Peak (%)	None[a] (%)	Walk-On Board (%)	Peak (%)	Before Peak (%)	After Peak (%)	None[a] (%)
10	–4	–7	7	7	–6	–3	–5	8	7	–5
20	–8	–13	14	14	–12	–7	–10	14	14	–10
30	–11	–19	19	19	–18	–10	–15	19	19	–15
40	–15	–24	23	22	–24	–14	–19	23	23	–20
50	–19	–29	26	25	–29	–17	–23	26	26	–24

[a]Includes those ferry riders who would make the trip some other way or not at all.

TABLE 7 Effects of Walk-On Ferry Fare Discounts

	Discretionary Trips					Nondiscretionary Trips				
		Drive-On Board					Drive-On Board			
Walk-on Fare Discounts (%)	Walk-On Board (%)	Peak (%)	Before Peak (%)	After Peak (%)	None[a] (%)	Walk-On Board (%)	Peak (%)	Before Peak (%)	After Peak (%)	None[a] (%)
10	2	0	0	0	0	2	0	0	0	0
20	4	–1	–1	–1	–1	3	–1	–1	–1	–1
30	5	–1	–1	–1	–1	5	–1	–1	–1	–1
40	6	–1	–1	–1	–2	6	–1	–1	–1	–1
50	7	–1	–1	–1	–2	7	–2	–2	–1	–1

[a]Includes those ferry riders who would make the trip some other way or not at all.

TABLE 8 Effects of Blended Ferry Fare Increases

	Discretionary Trips					Nondiscretionary Trips				
		Drive-On Board					Drive-On Board			
Blended Fare Increases (drive-on/walk-on) (%)	Walk-On Board (%)	Peak (%)	Before Peak (%)	After Peak (%)	None[a] (%)	Walk-On Board (%)	Peak (%)	Before Peak (%)	After Peak (%)	None[a] (%)
20/10	3	–3	–3	–3	6	3	–4	–4	–4	5
40/20	9	–11	–12	–11	19	9	–12	–13	–12	16
60/30	14	–22	–23	–22	33	15	–23	–24	–23	28
80/40	20	–35	–36	–35	47	21	–36	–37	–36	39
100/50	26	–50	–52	–51	60	27	–51	–53	–51	51

[a]Includes those ferry riders who would make the trip some other way or not at all.

reliable estimates based on comparisons with actual historic fare elasticities. The data that were collected and the models that were developed were designed specifically to address the study objectives, which focused on fare policies. The work was not designed to address the effects of changes in access–egress conditions or other types of changes in ferry service that might be considered in the future, such as provision of an onboard wireless local area network connection. However, similar stated preference survey methods could be used to address those issues. As the survey was administered primarily to peak drive-on customers, the model does not specifically account for a possible "rebound" response of current walk-on and off-peak passengers to improved peak crowding conditions. Similarly, the survey did not include travelers who are not now, but could become, ferry passengers as a result of service improvements. Again, a similar survey could be designed to address those types of responses.

This study provides a useful template for future work. Future studies conducted with this approach could address pricing and other service changes that are being considered at the time. The studies would ideally be coupled with continued data collection that identifies actual changes in passenger volumes as service conditions change.

ACKNOWLEDGMENTS

The authors acknowledge support from Washington State Department of Transportation Ferries Division, the Washington State Transportation Commission, and its consultants for this work.

REFERENCES

1. *Final Long-Range Plan.* Washington State Department of Transportation Ferries Division, Olympia, 2009.
2. Mayworm, P. D., A. M. Lago, and J. M. McEnroe. *Patronage Impacts of Changes in Transit Fares and Services.* Urban Mass Transportation Administration, U.S. Department of Transportation, Washington, D.C., 1980.
3. Fleishman, D., N. Shaw, A. Joshi, R. Freeze, and R. Oram. *TCRP Report 10: Fare Policies, Structures, and Technologies.* TRB, National Research Council, Washington, D.C., 1996.
4. Louviere, J. J., D. A. Hensher, and J. D. Swait. *Stated Choice Methods: Analysis and Applications.* Cambridge University Press, Cambridge, United Kingdom, 2000.
5. Ben Akiva, M., and S. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press, Cambridge, Mass., 1985.
6. Warburg, V., C. R. Bhat, and T. J. Adler. Modeling Demographic and Unobserved Heterogeneity in Air Passengers Sensitivity to Service Attributes in Itinerary Choice. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1951,* Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 7–16.
7. Axhausen, K. W., S. Hess, A. Koenig, J. Bates, M. Bierlaire, and G. Abay. State of the Art Estimates of the Swiss Value of Travel Time Savings. *Transport Policy,* in press.
8. Train, K. *A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit.* Working paper. Department of Economics, University of California, Berkeley, 2001.

The opinions expressed in, and the content of, this paper are the responsibility of the authors and not of the Washington State Department of Transportation or the others involved in this effort.

The Transportation Planning Applications Committee peer-reviewed this paper.

Changing Assignment Algorithms

Price of Better Convergence

Michael Florian and Shuguang He

The mainstay method of equilibrium assignment methods is based on adaptation of the linear approximation algorithm. Practically all commercial software packages for transportation planning offer a version of this algorithm. In the early days of personal computing, when random-access memory (RAM) was limited, this method was the most appropriate one to use because it requires little intermediate storage. As personal computers became more powerful and RAM became plentiful, the drawbacks of the linear approximation method became evident to practitioners. A measure of convergence is the relative gap, which measures the relative difference between total travel time and total travel time on the shortest paths. Relative gaps of less than 10^{-4} are difficult to reach with this method. Alternative assignment methods, based on algorithms that have better convergence rates, are known and can obtain finer solutions. Changing to new algorithms would appear to be a trivial task; however, it is not the case. The issues related to changing assignment algorithms pertain to the uniqueness of equilibrium paths, flows, and times. Examples of the expected changes in results for both standard multiclass assignments and for one complex model, which is equilibrated with feedback procedures, are presented. An adaptation of the projected gradient with path flows is used to represent the modern algorithms, which can reach relative gaps of 10^{-6} or better. Differences in relevant results are relatively small. Nevertheless, practitioners are careful to reproduce results and may face a challenge to accept slightly different results with a faster converging algorithm.

The mainstay of equilibrium assignment methods for a long time has been adaptation of the linear approximation algorithm of Frank and Wolfe (*1*). Practically all commercial software packages for transportation planning offer a version of this algorithm. In the early days of personal computing, when random-access memory (RAM) was limited, this method was the most appropriate one to use. However, as is well known, the linear approximation algorithm has the drawback that it requires a large number of iterations to obtain a very fine solution. A common measure of convergence is the relative gap, which measures the relative difference between total travel time and total travel time on the shortest path, at a given solution. Relative gaps of less than 10^{-4} are difficult to reach even if a very large number of iterations (~1,000) are carried out. Any path analysis requires rerunning the assignment or storing paths for a limited number of iterations.

INRO Consultants, 5160 Decarie Boulevard, Suite 610, Montreal, Quebec H3X 2H9, Canada. Corresponding author: M. Florian, mike@inro.ca.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 67–75.
DOI: 10.3141/2176-08

As personal computers became more powerful and RAM became plentiful, new opportunities became available for developing faster converging methods. Alternative methods implementing algorithms that have better convergence rates are available now in commercial software packages. Relative gaps on the order of 10^{-6} can now be obtained in reasonable computing times. Changing to the new algorithms would appear to be a trivial task; however, it is not the case. The issues related to changing algorithms pertain to uniqueness of equilibrium paths, flows, and times.

For a standard equilibrium assignment model with multiple classes, in which a single delay function is used for all the classes, the only quantities that are unique are the total link flows and the origin–destination (O-D) impedances. This fact is well known; see, for instance, Sheffi (*2*) and Rossi et al. (*3*). The class flows are not necessarily unique nor are any other attributes that depend on path analyses, such as tolls paid and distance traveled. The question then is how the results obtained with a current assignment will change when a new assignment method is adopted. Practitioners are generally very cautious to reproduce results. But, adopting a new assignment algorithm may change some results.

This paper provides insight into the expected changes in results for both standard multiclass assignments and for complex models that are equilibrated with feedback procedures. An adaptation of the projected gradient with path flows (*4*) is used to represent the modern algorithms, which produce very finely converged equilibrium assignments. Computing times are reduced drastically and changes in the most relevant results are slight. For other faster converging traffic assignment algorithms the interested reader may refer to Bar-Gera (*5*), Dial (*6*), and Slavin et al. (*7*).

MULTICLASS EQUILIBRIUM ASSIGNMENT MODEL AND ITS UNIQUENESS PROPERTIES

The most common mathematical formulation of the multiclass equilibrium assignment model is given next. Consider a transportation network model with several classes c, $c \in C$, of vehicular flow on the directed links of the network. The arcs a, $a \in A$, model the transportation links (streets, highways, etc.). Let I be the set of O-D pairs. Fixed O-D demands g_i^c, $i \in I$, give rise to link flows v_a^c, $a \in A$, $c \in C$, and the cost of traveling on a link is given by a user cost (travel time) function, $s_a(v_a)$, where v is the vector $(v_a = \sum_c v_a^c)_{a \in A}$ of link flows over the entire network. Cost functions model time delay on a link, or more general costs such as tolls and fuel consumption, and are assumed to be non-negative. K_i^c, $i \in I$, $c \in C$, are a set of directed paths connecting O-D pair i, and $K = \bigcup_{i,c} K_i^c$ is the set of all paths. The demand between O-D pairs uses the directed paths

K_i^c and the path flows h_k obey conservation of flow and non-negativity constraints:

$$\sum_{k \in K_i^c} h_k = g_i^c \qquad \forall i \in I \tag{1}$$

$$h_k \geq 0 \qquad \forall k \in K \tag{2}$$

Link flows are obtained by adding the corresponding path flows and are given by

$$v_a^c = \sum_{i=I} \sum_{k \in K_i^c} \delta_{ak} h_k \qquad \forall a \in A \tag{3}$$

where $\delta_{ak} = 1$ if link a belongs to path k and 0 otherwise.

Let Δ be the $|A| \times |K| \times |C|$ arc-path incidence matrix (δ_{ak}) so that $v = \Delta h$, where h is the vector $(h_k)_{k \in K}$ of path flows for all O-D pairs. The cost s_k^c $(= s_k(h), k \in K_i^c)$ for each path k and each class c is then defined by

$$s_k^c = \sum_{a \in A} \delta_{ak} s_a(\Delta h) \qquad \forall k \in K_i^c, i \in I, c \in C \tag{4}$$

and u_i^c $(= u_i(h), h \in K_i^c)$ is by definition the cost of the least-cost path for each class of traffic and each O-D pair i:

$$u_i^c = \min_{k \in K_i^c} s_k^c \qquad \forall i \in I, c \in C \tag{5}$$

The equilibrium traffic assignment is obtained by solving the convex cost optimization problem (*8*):

$$\min \sum_a \int_0^{v_a} s_a(y)dy$$

$$\sum_{k \in K_i^c} h_k = g_i^c \qquad \forall i \in I, c \in C$$

$$h_k \geq 0 \qquad \forall k \in K$$

$$v_a^c = \sum_{i \in I} \sum_{k \in K_i^c} \delta_{ak} h_k \qquad \forall a \in A, c \in C \tag{6}$$

The uniqueness properties of this model [see also Florian and Hearn (*9*)] are that the total link flows v_a, $a \in A$, are unique as well as the O-D travel times (costs) u_i^c, $i \in I$, $c \in C$. However, the class flows v_a^c, $a \in A$, $c \in C$, are not unique nor are the paths $k \in K_i^c$, $i \in I$, $c \in C$, or the path flows h_k, $k \in K$, because all classes share the same cost function (if each class has a separate and different cost function, the class flows v_a^c, $a \in A$, $c \in C$, are unique but the paths are not unique). This factor has some important implications for the results of a wide variety of analyses that rely on the analysis of paths such as "select link," "generalized select link," O-D matrix adjustment methods, and any analyses that use class flows in scenario comparisons.

Consider a planner who used the same code for the linear approximation method for 10 to 20 years. The nonuniqueness properties of the model would not be noticeable. Running the same software will produce the same results, even if the quantities computed are not unique. So, from this point of view, the uniqueness properties of the model are elusive. But, if one were to use different implementations of the linear approximation method, as it is offered in different software packages, the nonuniqueness properties would be visible. The results will be similar but not exactly the same for the quantities that are not unique. If one continues to use the same software package, some of the issues raised by the uniqueness properties are not important. However, if one were to change assignment algorithms to any one of the more recent methods that exhibit better convergence properties than the linear approximation method, the difference in results would be noticeable.

If an assignment algorithm produces slightly different class flows for a multiclass assignment model or if the "select link" results are slightly different than those obtained with the linear approximation method, it does not mean that the new results are wrong in any sense. They are simply different because of the model properties. The old results obtained with the linear approximation method are not more correct in any sense. But obtaining different results may be difficult to interpret by some practitioners who are not concerned with the uniqueness properties of the model.

COMPARING A FASTER ASSIGNMENT ALGORITHM WITH LINEAR APPROXIMATION METHOD

To compare results obtained with the linear approximation algorithm with one of the more recent algorithms that exhibits faster convergence and can produce much finer solutions, an adaptation of the projected gradient (*2*) is used with path flows to represent the modern algorithms. A cursory description of the algorithm is as follows:

1. Compute the average cost of all used paths (by O-D pair).
2. Reduce the flow of paths that have a larger cost than the average.
3. Increase the flow on paths that have a smaller cost than the average.
4. Just keep the paths with positive flow.
5. Add a path if it is shorter than the current equilibrated solution.

There is, of course, the need to compute a step size to determine the magnitude of the flow change. Its performance, compared with the linear approximation method, is shown for a few large-scale single and multiclass equilibrium assignments that originate in practice.

For each chosen data set the flows obtained after 1,000 iterations of the linear approximation method were compared with flows computed with the projected gradient method, which exhibited a relative gap of 10^{-6}. [The relative gap is defined as the ratio of (total travel time − total travel time on shortest paths) and the total travel time.]

The Montreal, Quebec, Canada, regional planning network, shown in Figure 1, has 1,425 zones, 13,491 nodes, and 33,540 links. Three classes of traffic are modeled: private car, light trucks, and heavy trucks. The linear approximation method barely gets a relative gap of less than 10^{-4} while the projected gradient method reaches a relative gap of 10^{-6} in a reasonable time (see Figure 2). The Sydney, Australia, network shown in Figure 3 is one of the planning networks used by the Road and Traffic Authority of New South Wales, Australia. It consists of 1,155 centroids, 12,893 regular nodes, 34,551 directional links, and 8,415 turntable entries for one class of traffic. The improvement in convergence offered by the projected gradient method (Figure 4) is more pronounced even though the linear approximation method reaches a relative gap of 10^{-3} quicker. Figure 5 shows the 2000 Base South Corridor regional planning models of Metro Portland, Oregon. It contains 1,260 zones, 8,794 nodes, 26,091 links, and 7,010 turns, and four classes of traffic are assigned: single-occupancy vehicle, high-occupancy vehicle, heavy trucks, and medium trucks. Improvement

FIGURE 1 Montreal network.

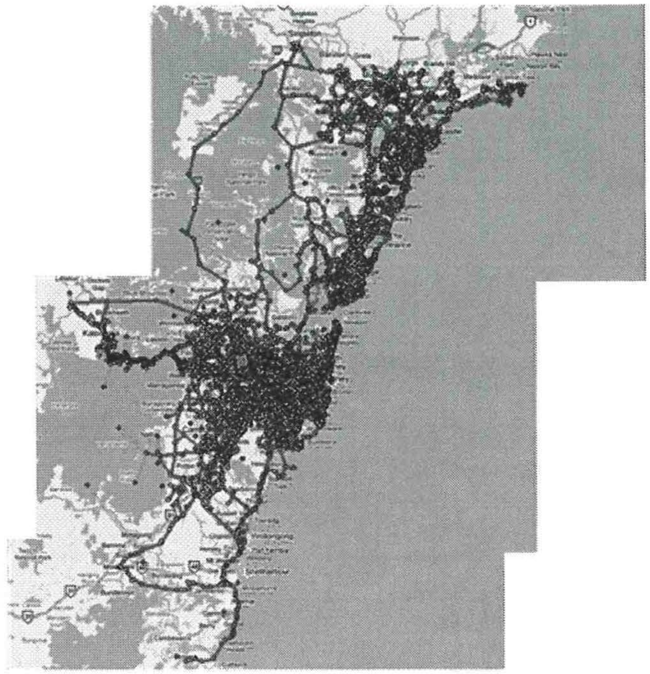


FIGURE 3 Network of Sydney Road and Traffic Authority of New South Wales.

in assignment results and running times is similar to that in the other networks discussed above (see Figure 6). The last example is a network of Phoenix, Arizona, in Figure 7 that was used by the Maricopa Association of Governments for planning applications. It contains 2,041 centroids, 12,938 regular nodes, 39,731 directional links, and 1,896 turntable entries. The results displayed in Figure 8 correspond to the assignment of a single class of traffic. The improvements in convergence are evident as well for this network.

ILLUSTRATION OF NONUNIQUE RESULTS OF EQUILIBRIUM ASSIGNMENTS

The most relevant questions concern how different the nonunique results are if one compares the results of the linear approximation method with those obtained with the projected gradient method. Several such differences are presented.

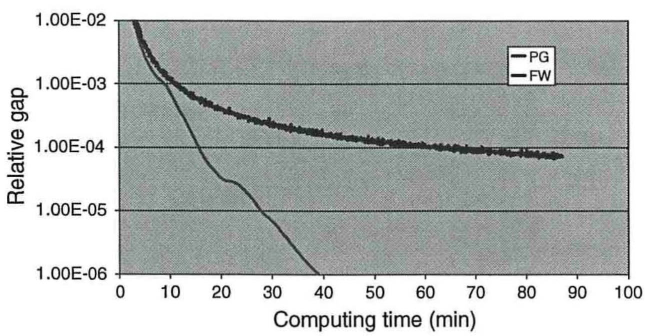


FIGURE 2 Convergence comparison for Montreal network (PG = projected gradient method, FW = linear approximation method of Frank and Wolfe).

The first example is a "select link" analysis of one of the links of the Chicago, Illinois, test network (a network publicly available for testing assignment algorithms). The network displayed in Figure 9 contains 1,790 zones 11,192 regular nodes, and 39,018 directional links. A single class of traffic is assigned. Figures 10 and 11 show the selected flows that use the link that has the largest flow value obtained with the linear approximation method and with the projected gradient, respectively. The differences are minimal—on the order of one to three trips on the largest flow links, which have a flow on the order of 1,000. Nevertheless, the O-D pairs that contribute the flows to the selected links are different as the paths are not unique.

Next, the nonuniqueness of class flows is exhibited by comparing the two assignment methods on the Montreal network. The linear approximation assignment was run to a relative gap of 10^{-4} and the projected gradient assignment was run to a relative gap of 10^{-6}. A comparison of the total flows, which are unique, is shown in Figure 12. The private car flows, which are shown in Figure 13, are very similar. Some differences can be observed in the comparison of light and heavy truck flows in Figures 14 and 15, respectively. The general pattern is the same, but the nonunique properties of the class flows are

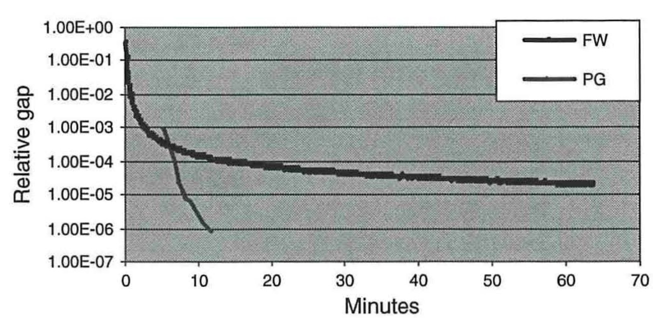


FIGURE 4 Convergence comparison for Sydney network.

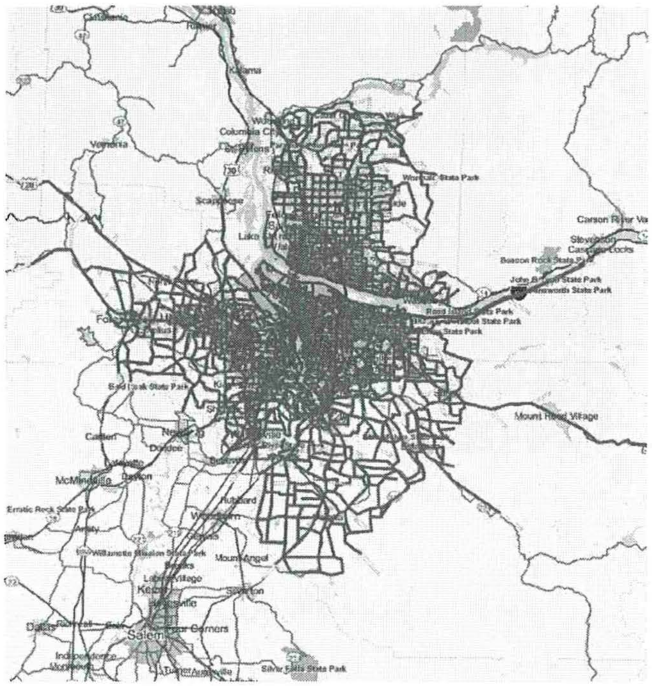


FIGURE 5 Metro Portland network.

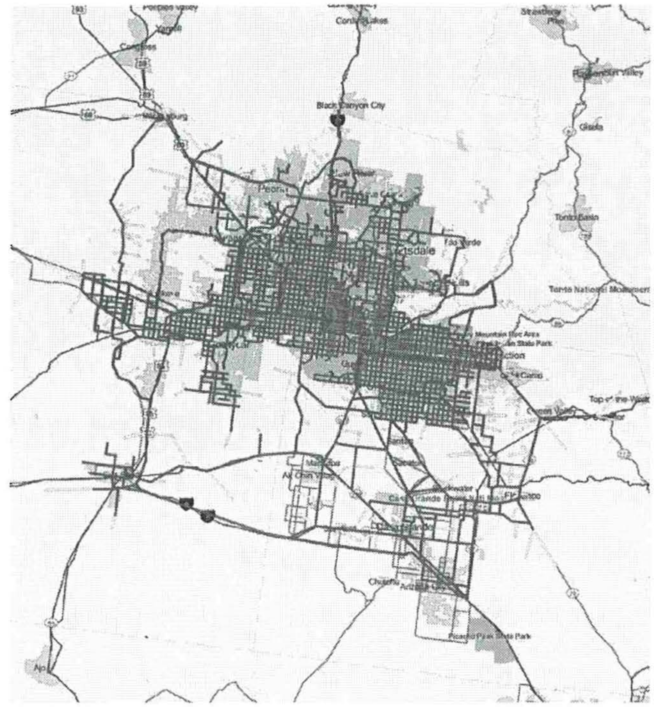


FIGURE 7 Network of Phoenix Maricopa Association of Governments.

more evident in the last two comparisons. The linear regression of the truck flows computed with the linear approximation flows versus the projected gradient flows yield an r^2 of 0.9999.

The other question of interest is how the results of a complex planning model, which uses many submodels for demand forecasting, change when the linear approximation method is replaced by the projected gradient method. Such changes are demonstrated with the planning model of the Puget Sound Regional Council (PSRC). The transportation network is shown in Figure 16 and the model structure is presented in block diagram in Figure 17 (*10*). The only change made to the entire process was to replace the assignment, which is indicated in Figure 17 as the last box in the second column.

The PSRC transportation planning model tested here and shown in Figure 17 contains 15 modes, 30 transit vehicle types, 1,155 zones, 834 transit lines, 5,888 regular nodes, 25,856 transit line segments, 20,633 directional links, and 16,864 turntable entries. Eleven classes of traffic are assigned. Figure 18 compares the results obtained with the linear approximation assignment (first column of numbers) with the results obtained with the projected gradient assignment (second column of numbers). The total execution time was reduced by 50%, from 20.4 to 10 h, with minimal changes in the results. Because the O-D travel generalized costs are unique for each class, the feedback equilibration process used travel impedances, even before replacement of the assignment, in an averaging scheme. The results are not identical but they are sufficiently close to justify the change of assignment methods used.

CONCLUSIONS

The faster converging assignment methods are presented here. Changing the assignment algorithms leads to some changes in the results due to the nonuniqueness properties of the deterministic multiclass network equilibrium model. In the future, perhaps the issue of nonuniqueness will be resolved with new developments in assignment methodology. Until then, practitioners may choose to keep the linear approximation algorithm or change to faster converging methods with some change in the results obtained.

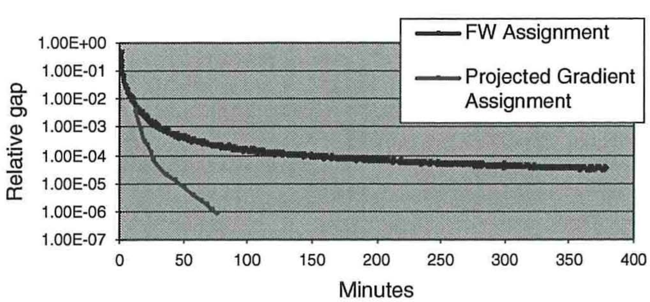


FIGURE 6 Convergence comparison for Metro Portland network.

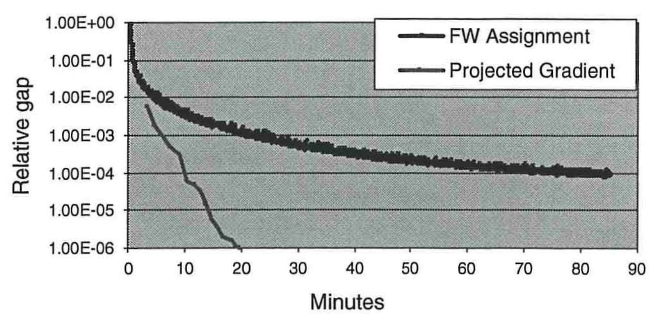


FIGURE 8 Convergence comparison for network of Phoenix Maricopa Association of Governments.

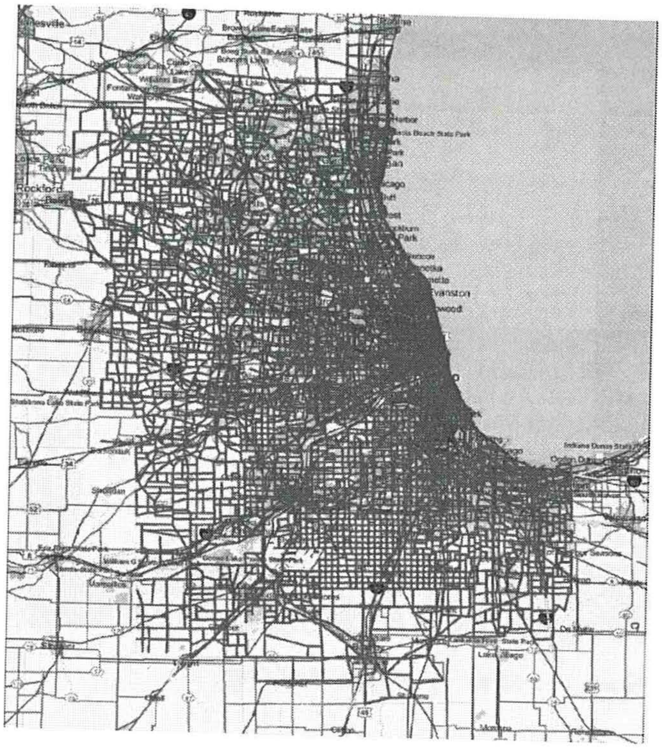


FIGURE 9 Chicago test network.

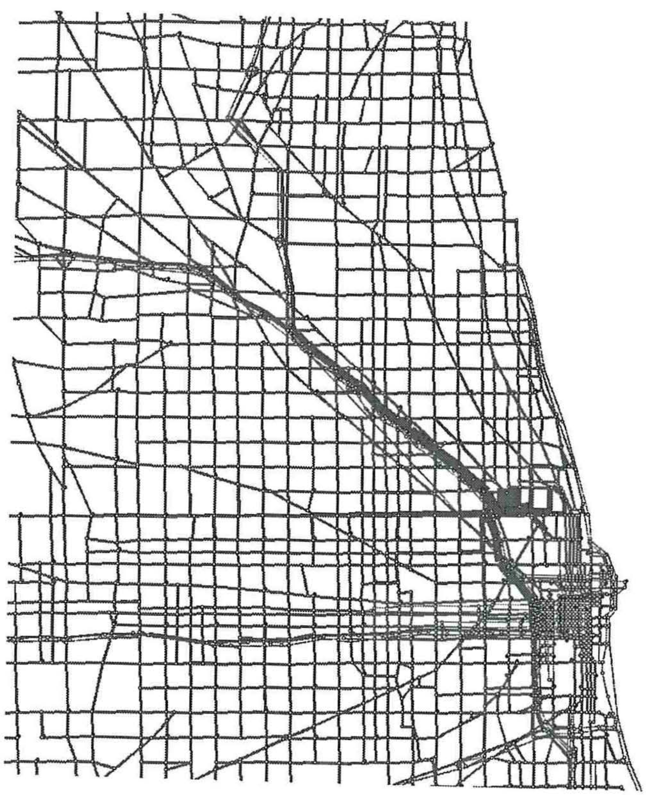


FIGURE 10 Select link linear approximation run to relative gap of 10^{-4}.

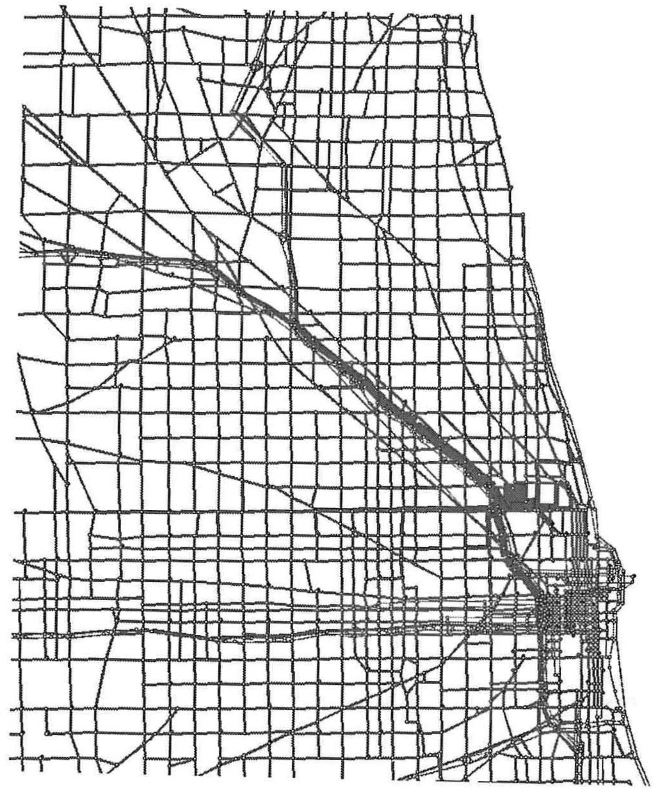


FIGURE 11 Select link projected gradient run to relative gap of 10^{-6}.

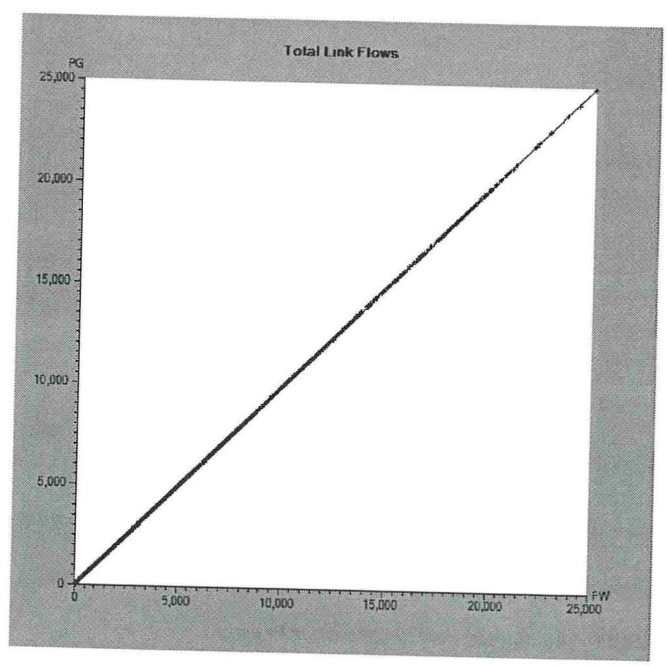


FIGURE 12 Comparison of total flows.

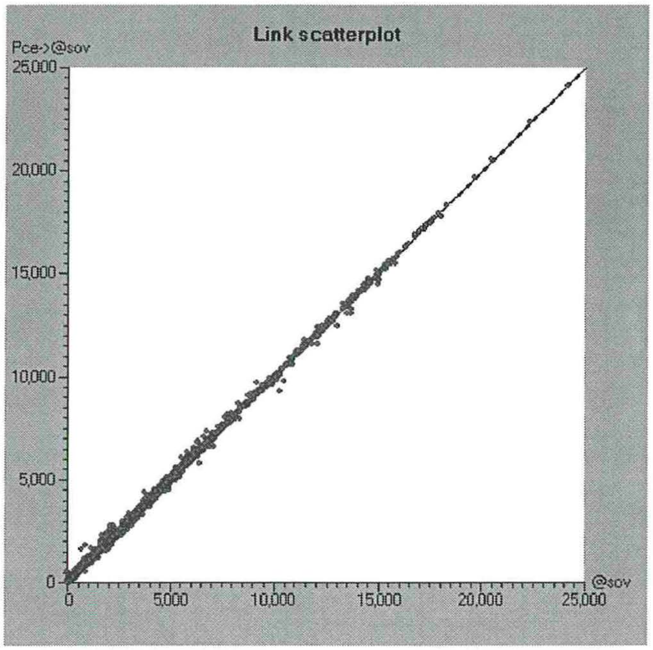


FIGURE 13 Comparison of single-occupancy flows.

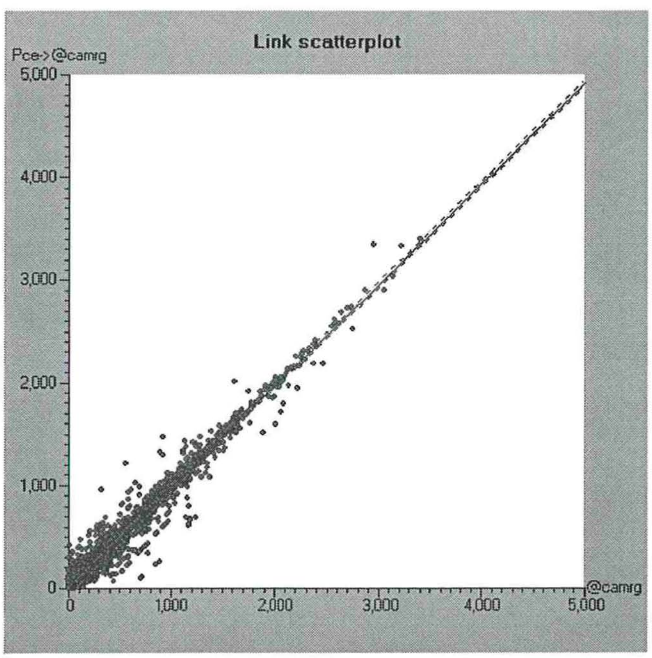


FIGURE 14 Comparison of light-truck flows.

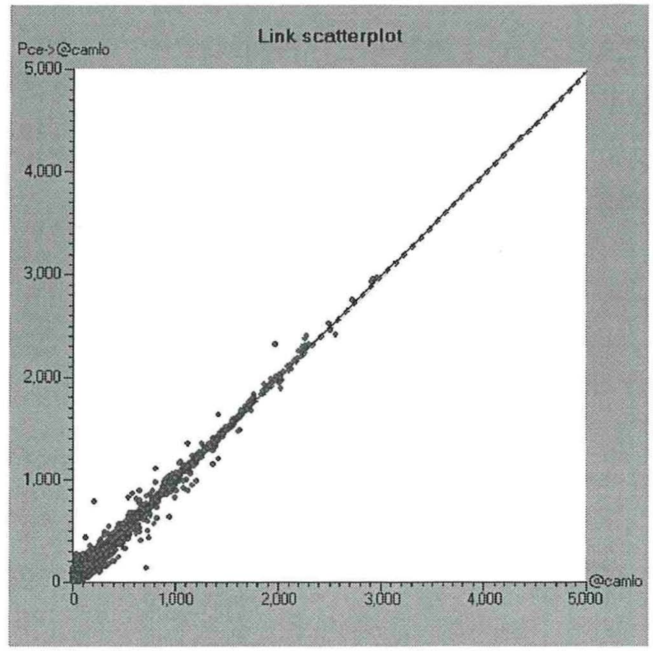


FIGURE 15 Comparison of heavy-truck flows.

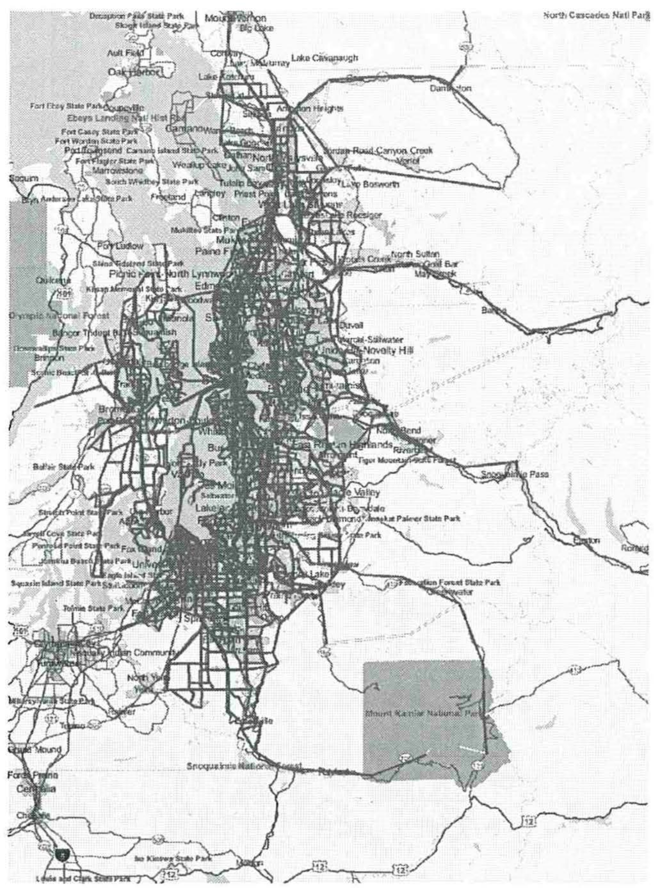


FIGURE 16 PSRC regional network.

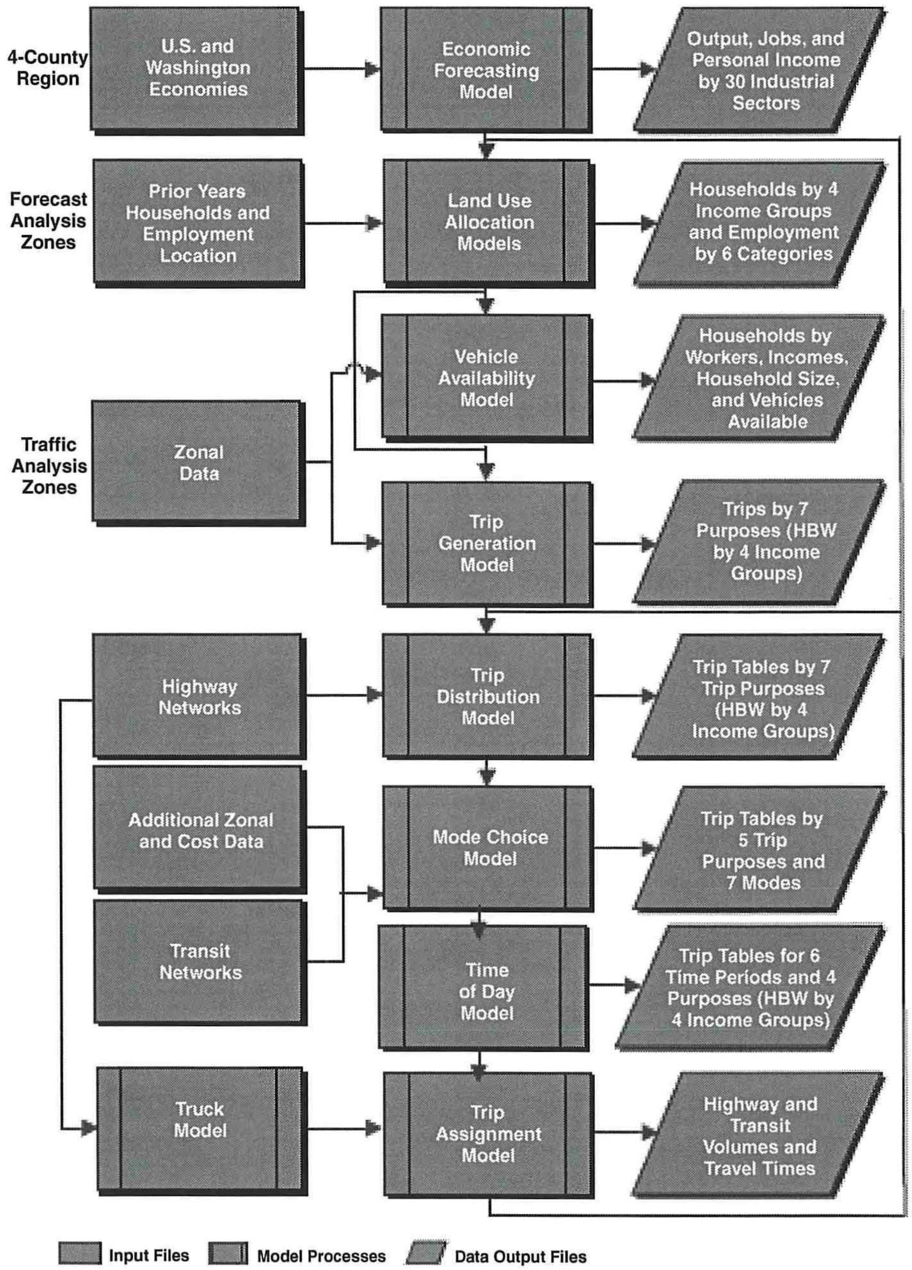


FIGURE 17 PSRC land use and travel demand forecasting process (*10*).

		2000_v1.0bb run time = 20.4 hr 5/14/2008	2000_v1.0bb 5.25 test run time = ~10 hr 10/15/2008	Difference	% Diff
VMT	AM	14,029,092	14,003,212	-25,880	-0.2%
	MD	27,977,890	28,059,674	81,784	0.3%
	PM	17,938,798	17,996,364	57,566	0.3%
	EV	13,049,526	12,978,556	-70,970	-0.5%
	NT	6,808,230	6,800,137	-8,093	-0.1%
	Daily	79,803,536	79,837,943	34,407	0.0%
VMT	freeway	35,386,219	35,415,529	29,311	0.1%
	arterial	34,360,506	34,366,080	5,574	0.0%
	connector	10,056,764	10,056,297	-467	0.0%
	TOTAL	79,803,489	79,837,906	34,417	0.0%
VHT	AM	421,798	421,724	-75	0.0%
	MD	827,716	831,812	4,095	0.5%
	PM	623,837	631,438	7,601	1.2%
	EV	374,600	371,981	-2,619	-0.7%
	NT	164,306	164,155	-151	-0.1%
	Daily	2,412,257	2,421,109	8,852	0.4%
VHT	freeways	790,386	796,465	6,080	0.8%
	arterials	1,022,010	1,024,810	2,800	0.3%
	connectors	599,864	599,832	-32	0.0%
	TOTAL	2,412,260	2,421,107	8,847	0.4%
Avg Speed	AM	33.3	33.2	-0.1	-0.2%
	MD	33.8	33.7	-0.1	-0.2%
	PM	28.8	28.5	-0.3	-0.9%
	EV	34.8	34.9	0.1	0.2%
	NT	41.4	41.4	0.0	0.0%
	Daily	33.1	33.0	-0.1	-0.3%
Avg Speed	freeways	44.8	44.5	-0.3	-0.7%
	arterials	33.6	33.5	-0.1	-0.3%
	connectors	16.8	16.8	0.0	0.0%
	TOTAL	33.1	33.0	-0.1	-0.3%
Person Trips (All Trips - no school)	work/col	2,086,006	2,086,006	-1	0.0%
	non-work	9,869,872	9,869,873	1	0.0%
	TOTAL	11,955,878	11,955,878	0	0.0%
Mode Shares (HBW - no college)	SOV	80.1%	80.3%	0.2%	0.2%
	carpool	7.2%	7.2%	0.0%	0.0%
	transit	8.1%	8.0%	-0.2%	-2.0%
	- Transit-walk	6.6%	6.6%	0.0%	-0.2%
	- Transit-auto	1.5%	1.3%	-0.1%	-9.9%
	bike	1.6%	1.6%	0.0%	0.0%
	walk	3.0%	3.0%	0.0%	-0.1%
Mode Shares (Non Work)	SOV	45.9%	45.9%	0.0%	0.0%
	carpool	44.5%	44.5%	0.0%	0.0%
	transit	1.8%	1.8%	0.0%	0.1%
	bike	1.0%	1.0%	0.0%	0.0%
	walk	6.8%	6.8%	0.0%	0.0%

FIGURE 18 Comparison of aggregate results: linear approximation versus projected gradient assignments (VMT = vehicle miles traveled, VHT = vehicle hours traveled, HBW = home-based work, AM = morning, MD = midday, PM = afternoon, EV = evening, NT = night, SOV = single-occupancy vehicle). (Source: Computations by PSRC staff.)

ACKNOWLEDGMENTS

The authors express their appreciation to the Ministry of Transportation of Quebec, Metro Portland, Sydney Road and Traffic Authority, Maricopa Association of Governments, and PSRC for granting permission to use their data in the preparation of this document.

REFERENCES

1. Frank, M., and P. Wolfe. An Algorithm for Quadratic Programming. *Naval Research Logistics Quarterly,* Vol. 3, 1956, pp. 95–110.
2. Sheffi, Y. *Urban Transportation Networks: Equilibrium Analysis and Mathematical Programming Methods.* Prentice Hall, Englewood Cliffs, N.J., 1985.
3. Rossi, T. F., S. McNeil, and C. Hendrickson. Entropy Model for Consistent Impact-Fee Assessment. *Journal of Urban Planning and Development,* Vol. 115, 1989, pp. 51–63.
4. Florian, M. A., I. Constantin, and D. Florian. A New Look at Projected Gradient Method for Equilibrium Assignment. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2090,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 10–16.
5. Bar-Gera, H. Origin-Based Algorithm for the Traffic Assignment Problem. *Transportation Science,* Vol. 36, 2002, pp. 398–417.
6. Dial, R. A Path-Based User-Equilibrium Traffic Assignment Algorithm That Obviates Path Storage and Enumeration. *Transportation Research B,* Vol. 40, 2006, pp. 917–936.
7. Slavin, H., J. Brandon, and A. Rabinowicz. An Empirical Comparison of Alternative User Equilibrium Traffic Assignment Methods. Presented at European Transport Conference, Strasbourg, France, 2006.
8. Beckman, M., C. McGuire, and C. Winsten. *Studies in the Economics of Transportation.* Yale University Press, New Haven, Conn., 1956.
9. Florian, M., and D. Hearn. Network Equilibrium Models and Algorithms. In *Handbooks in Operations Research and Management Science, Vol. 8* (M. O. Ball, T. L. Magnanti, C. L. Monma, and G. L. Nemhauser, eds.), North-Holland, Amsterdam, Netherlands, 1995, pp. 485–550.
10. *PSRC Travel Model Documentation (for Version 1.0) Updated for Congestion Relief Analysis.* Final report. Cambridge Systematics, Cambridge, Mass., 2007. www.psrc.org/assets/1511/model_doc_final_.pdf.

The Transportation Planning Applications Committee peer-reviewed this paper.

Intelligent Transportation System as Evaluation Tool in a Regional Demand Modeling Environment

Implementation in Florida Standard Urban Transportation Model Structure

Yan Xiao, Mohammed Hadi, Halit Ozen, and Vidya Mysore

A number of tools have been developed to support the evaluation of intelligent transportation system (ITS) alternatives. However, the methodologies and some associated parameters in these tools were developed and selected before widespread deployments of ITSs. Significant experience gained with ITS in recent years warrants a new assessment of the evaluation methodologies and parameters of these systems. There are also advantages to incorporating the evaluation of ITS deployments as part of existing regional demand forecasting modeling environments rather than using external tools to perform this functionality. This paper discusses the development and implementation of a tool and methodologies to estimate the benefits and costs of these systems as part of a travel demand forecasting modeling environment. It also presents an application of the developed tool to evaluate two of the most widely deployed types of ITS: incident management and advanced traveler information systems. Case study results indicate that the methodologies developed in this study can be used to assess ITS deployments.

The planning of intelligent transportation systems (ITSs) requires detailed evaluation of the performance and costs of ITS deployment alternatives relative to each other and to other types of transportation system improvement alternatives. A number of sketch planning tools have been developed to support the evaluation of ITS alternatives. These tools range in details from a simple spreadsheet with simplified assumptions such as the screening analysis for ITS (SCRITS) (*1*) to more sophisticated tools like the ITS deployment analysis system (IDAS) (*2*). Despite the modeling capabilities of these tools, most evaluation methodologies and some parameters were established when ITS was just starting in the 1990s. The ITS field has experienced considerable developments and advancements since then. Thus, a fresh look is needed of the evaluation methodologies used in sketch planning tools based on what has been learned in the past 10 years of ITS deployments.

Y. Xiao and H. Ozen, EC 3730, and M. Hadi, EC 3605, Department of Civil and Environmental Engineering, Florida International University, 10555 West Flagler Street, Miami, FL 33174. V. Mysore, Systems Planning Office, Florida Department of Transportation, 605 Suwannee Street, MS 19, Tallahassee, FL 32399. Corresponding author: M. Hadi, hadim@fiu.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 76–83.
DOI: 10.3141/2176-09

There are also advantages to incorporating the evaluation of ITS deployments as part of existing regional demand forecasting modeling environments rather than using external tools to perform this functionality. The existing tools include internal models that differ from the calibrated regional demand models. This difference results in inconsistencies in the evaluation and forecasting processes between the evaluation tool and the regional models. A Northeastern Illinois case study, conducted by the Chicago Area Transportation Study (CATS) to evaluate IDAS capabilities (*3*), suggested that ITS evaluations should be incorporated as part of the CATS regional travel demand models. It was stated that such implementations will ensure the consistency of reporting measures such as better estimation of emissions, reducing the duplicated effort on converting the network files and demand matrices into the format required, and enhanced analysis capabilities.

The advancements in transportation demand forecasting models and the integration of these models with geographic information systems make them attractive environments for the development of ITS evaluation tools. The Florida Standard Urban Transportation Model Structure (FSUTMS) represents a formal set of modeling steps, procedures, software, file formats, and guidelines established by the Florida Department of Transportation (FDOT) for use in travel demand forecasting throughout the state of Florida. The software environment that is used as the software engine for the FSUTMS has powerful data handling and modeling capabilities that allow the incorporation of advanced evaluation of ITS deployments.

This paper reports on an effort initiated by FDOT to implement ITS evaluation as part of the FSUTMS modeling environment. The paper also presents an application of the developed tool to evaluate two of the most widely deployed types of ITS: incident management (IM) and advanced traveler information systems (ATIS).

OVERVIEW OF DEVELOPED TOOL

The developed ITS evaluation tool was implemented as part of the FSUTMS demand modeling environment by using the script language of Cube, the modeling engine of the FSUTMS. Two stakeholder workshops were conducted as part of the tool development effort. The first workshop was conducted before development to identify and confirm the development requirements. The second workshop was conducted toward the end of development to confirm that all requirements had been met. The identified requirements

addressed the types of ITS deployments to be evaluated and assessed impacts and performance measures, supporting modules, and other general evaluation requirements.

The types of ITS deployments that can be evaluated with the current version of the developed tool include ramp metering, IM systems, highway advisory radio and dynamic message signs (DMS), ATIS, managed lanes, signal control, transit vehicle signal priority, emergency vehicle signal priority, advanced public transit systems, smart work zones, and road weather information systems. The study team conducted an extensive review and assessment of the methods and procedures used in previous studies and existing sketch planning tools to evaluate ITS deployments. On the basis of this review and assessment, methods were identified and implemented to evaluate each of the chosen ITS deployment options. Additional evaluation procedures can be added to assess other ITS deployment types that are not currently included in the evaluated ITS deployment types.

The ITS evaluation methods require three types of parameters: factors that reflect the impacts of ITS technologies and strategies on various performance measures, parameters to identify the initial and recurrent costs of ITS deployments, and dollar values to convert ITS deployment benefits such as the reductions in delays and crashes to monetary values. The default values for these parameters were selected in this study based on an extensive review and assessment of the information available on the subject and surveys of ITS engineers in Florida. Analysts can change these parameters if better local values are available for a given analysis.

Depending on the types of the evaluated ITS deployments, the tool can produce various performance measures including vehicle miles of travel, vehicle hours of travel, average speed, number of accidents by severity level (fatality, injury, and property damage only), fuel consumption (gallons), monetary benefits to users and agency, and emission level by pollutant (hydrocarbon, carbon monoxide, and oxides of nitrogen). Modules were identified and implemented to estimate these measures as part of the developed tool.

The tool and methods developed in this study are being implemented or considered for implementation in a number of Florida metropolitan planning organization–approved regional demand models.

BACKGROUND OF IM AND ATIS EVALUATIONS

This paper presents the evaluation procedures of IM and ATIS deployments as examples of the developed methods and their applications. This section reviews how IM and traveler information systems have been evaluated in previous efforts.

IM Evaluation

IM aims at coordinating the activities of transportation agencies, police, and emergency services; facilitating incident detection, verification, response, and clearance; and therefore reducing incident duration and minimizing the negative impacts of incidents. In the absence of field measurements of the reduced delay due to IM systems, one of three methods can be used to estimate the reduction in delay based on the reduction in incident duration: queuing analysis, shockwave analysis, and simulation analysis. Traffic simulation analysis is a powerful method to analyze the benefits of traffic management and IM. However, the use of simulation models is expensive in terms of data collection requirements, model input preparation, and calibration.

When comparing queuing and shockwave analysis, queuing analysis is by far the most widely used method to identify incident impacts with and without IM strategies. Detailed discussions of deterministic queuing analysis and shockwave analysis can be found in traffic flow theory textbooks (*4*). A study by Rakha and Zhang (*5*) has demonstrated the consistency in delay estimates that are derived from deterministic queuing theory and shockwave analyses. They indicated that queuing theory provides a simple and accurate technique for estimating delay at highway bottlenecks.

Although not stated in the IDAS user manual, IDAS actually uses queuing theory in calculating incident delays with assumptions made about the rate of lane blockage incidents per million vehicle miles, incident duration, and peaking characteristics of traffic within the analysis period. IDAS applied a default of 9% reduction in incident duration for incident detection and verification systems, 39% for incident response and management systems, and 51% for the combination (*2*). In addition to the mobility benefits, IDAS also assumes a shift of 10% to 21% of freeway fatalities to injuries, accompanied by a 15% to 42% reduction in fuel consumption and pollutant emissions with implementation of IM systems. The analysis of congestion management alternatives of the Kansas City, Missouri, region using IDAS (*6*) showed that deployment of the IM system produced a 7.2% reduction in incident delay in this region and a 40% reduction within the study area.

DMS and ATIS Evaluation

ATIS involve collecting, aggregating, and disseminating information to assist surface transportation travelers in moving from their origins to destinations. This information could include travel time, incident, roadway construction, weather, emergency, transit, optimal routes to destinations, and traveler service information. Information dissemination involves using a number of technologies for disseminating traveler information, such as highway DMS, highway advisory radio, traveler information telephone systems, and websites.

In existing sketch planning tools, a common way to calculate the benefits of DMS and ATIS is to multiply the number of trips that can access the traveler information system, the percentage of people using such information, the percentage of those people saving time, and the average delay saving per user benefiting from the system. These tools require the user to input the above parameters. For example, the SCRITS program applies the defaults of 1-h activation per day, 20% of drivers benefiting from DMS, and 3-min savings per driver for each DMS activation. In IDAS, the default values for DMS evaluation are 10% of the total time the message sign is turned on, 28% of vehicles passing the message sign and saving time, and average time savings of 11 min per traveler.

For web-based or phone-based ATIS, IDAS assumes that the market penetration is 0.5% for both types of ATIS in the year 2000. It further assumes that this percentage increases by 1% or 2% every year for the web-based ATIS but not for the phone-based ATIS. The corresponding percentage of time savings is a decreasing function of market penetration, ranging from 20% to 0%. Users can change these default values. A number of studies have applied the IDAS software to evaluate the benefits and costs of ATIS. The Ohio–Kentucky–Indiana regional council of governments performed such an analysis (*7*). In the study, the percentage of users benefiting from the telephone and web information services was changed from the default of 1% to 0.42% according to local conditions. A study by the Michigan Department of Transportation (*8*) recommended increasing the percentage of benefiting users from the telephone and web information

services from 1% to 1.4% and the delay saving from 15% to 20%. On the basis of a comprehensive literature review, Kristof et al. (*9*) recommended using IDAS to evaluate ATIS projects in Washington State. However, the study recommended that the percentage of benefiting users for a telephone-based information system should be increased from 0.5% to 10%. For the web-based information system, Kristof et al. (*9*) recommended changing the IDAS defaults to 3% of benefiting users for the year 2000, 8% for 2005, 13% for 2010, 20% for 2015, and 30% for 2020. The maximum amount of time saving was recommended to be 15% instead of the default of 20% for 10% market penetration and 5% instead of 0% for 60% market penetration.

Market Penetration and Compliance Rate

Market penetration and compliance rate are two critical parameters in the evaluation of ATIS. Market penetration is the percentage of travelers who access the information, whereas compliance rate is the percentage of those travelers who make changes in trip decisions based on the information.

The market penetration for DMS can be assumed to be all the vehicles passing the sign. For phone-based and web-based ATIS, there is a need to determine the proportions of travelers accessing the systems. Yim and Miller (*10*) evaluated the TravInfo regional ATIS in the San Francisco Bay Area, California, from September 1996 to September 1998. The results showed that only 9% of households were aware of this system and very few of them had ever tried it. Yim et al. (*11*) updated the study in 2002 and indicated that 4% of household survey respondents were using the Internet, 18% were using telephone as a source of pretrip information, and 2% were using a cell phone for en route information. A study by Peirce and Lappin (*12*) of the Seattle ATIS indicated that travelers acquire information on only about 10% of their trips and they change travel plans during 9% of these acquiring trips. This information indicates that only 0.09% of travelers may benefit from the system. The 511 Deployment Coalition (*13*) indicated that the number of calls per day could range from 1.5% to 2.5% of the region's population, depending on the geographic conditions of the area, the configuration of the transportation network, and congestion in the area.

To determine the benefits, it is also necessary to understand the diversion behavior of travelers (the compliance rate). Several researchers have used the stated preference approach in an attempt to determine the percentage of travelers changing trip decisions in response to information disseminated by ATIS technologies. The studies concluded, based on this type of survey, that the disseminated information can result in up to 60% to 70% of freeway traffic exiting the freeway ahead of a bottleneck, such as an incident location (*14–17*). However, information about the actual diversion due to traveler information has been limited. Several European field studies have found that DMS compliance rates range between 27% and 44% (*18*). It is expected, however, that the compliance rate in the case of phone-based and web-based ATIS, particularly during lane blockage conditions, is much higher than this rate for DMS because travelers calling 511 or accessing the web are more likely to modify their behaviors to save time.

Several studies also relate the diversion rate to delay reductions. Mahmassani (*19*) and Liu and Mahmassani (*20*) assumed that travelers switch routes with information based on a "boundedly rational" switching behavior with travelers not changing their routes as long as the difference between travel time on the subject route and alternative route is below a certain threshold—for example, a mean of 1 min or 20% of travel time, determined empirically from user behavior studies. Peeta et al. (*21*) investigated the impacts of DMS information content and other relevant factors on diversion rates using the stated preference method. The results indicated that 53% of drivers would divert to an alternative route when the expected delay on the current route exceeds 10 min. On the basis of a stated preference survey, Khattak et al. (*22*) found that 42.9% of respondents would definitely take alternative routes under jammed conditions. Huchingson and Dudek (*23*) developed a linear relationship between diversion rate and posted incident delays on the DMS, with zero diversion for zero delay and 95% diversion for 1-h delay.

METHODOLOGY

This section describes the methodology used to evaluate the impacts and costs of the combined IM and ATIS as part of the FSUTMS framework. The main impact of IM systems is to reduce travel times of travelers by reducing incident durations and thus the associated delays to travelers. In general, these systems are not expected to affect trip generation, trip distribution, mode split, or route choice. ATIS and DMS provide the real-time traffic information to travelers, allowing them to make better trip choices, particularly during incident conditions. The main impact of these systems is to allow travelers to change their routes during incident conditions. With ATIS, travelers may decide to change travel mode, change trip time, or cancel a trip under these conditions. In general, these systems are not considered to affect trip generation, trip distribution, model split, and traffic assignment during "normal" conditions.

Incident Management

Queuing analysis is used in this study to calculate incident delays, the major impacts of IM. According to queuing theory, the total delay (TD) for a given type of incident TD_i can be expressed as follows:

$$\mathrm{TD}_i = \frac{t_R^2(\mu-\mu_R)(\lambda-\mu_R)}{2(\mu-\lambda)} \tag{1}$$

where

t_R = incident duration, which varies with incident type;
λ = mean arrival rate;
μ = mean capacity under the prevailing condition; and
μ_R = capacity during the incident.

Considering the incident frequency, total incident delay (TAD) is expressed as the summation of delays due to all types of incidents:

$$\mathrm{TAD} = \sum_i (\mathrm{IN}_i \times \mathrm{TD}_i) \tag{2}$$

where IN_i dictates the number of incidents for incident type i, which is a function of average incident rate and vehicle miles traveled on the roadway segments. Implementation of IM reduces the incident duration and thus the total incident delays as indicated in Equation 1.

The relationships presented above were derived for freeway segments. Little work has been done to estimate the impacts of incidents on arterial delays. Yang et al. (*24*) found that the delay due to an incident on a signalized arterial street is 15% to 34% higher when there are signals on the highway segment compared with the seg-

ments without signals (uninterrupted flow segments). In addition, the incident frequency and average duration are expected to differ on arterials compared with freeways. To estimate the arterial incident delays, the same procedures as mentioned above are applied; however, the arterial incident rate and duration are used, and an adjustment factor is applied to the calculated TD to account for the additional incident delay on signalized arterial streets compared with uninterrupted flow highways.

$$\mathrm{TD}_{i,s} = a_s \mathrm{TD}_{i,u} \quad (3)$$

where subscript u represents uninterrupted traffic and s denotes signalized flow. The parameter a_s is assumed to be 1.25, based on the conclusions of Yang et al. (*24*).

The values of incident frequency, duration, lane blockage statistics, and volume variations during the analysis period are not built in the model. They are explicitly input by the users. Default values were identified for these parameters based on data from freeway corridors in Broward County, Florida, maintained by the FDOT District 4 ITS program (*25*), and incident statistics for an urban arterial street (US-1) in Miami–Dade County, Florida, maintained by the FDOT District 4 ITS program (*26*).

Improvement in safety is another important benefit of the application of IM systems. It is assumed that 21% of fatalities are shifted to injuries due to the quick incident detection, verification, and response of IM systems (*2*). However, in addition to the above benefits, a reduction in accident rate is assumed as the IM system reduces the period of time hazardous driving conditions exist due to primary incidents. In San Antonio, Texas, the crash rate decreased by 2.8% due to implementation of IM. Thus, an additional reduction factor of 2.8% was used in this study for fatal, injury, and property-damage-only accidents (*27*).

Instead of using default reduction factors for emissions and fuel consumption due to IM, the FSUTMS implementation calculates the emission and fuel consumption with and without IM based on the speeds of queued and nonqueued vehicles and the vehicle miles in queue. Outside the queue, speed is calculated by using the relationship between speed and volume–capacity ratios, calibrated for the traffic demand forecasting models of the region. With incidents, the average queue length is obtained by using the queuing equations. The speed within the queue is assumed to be 2.5 mph. The average queue length as produced by the queuing equations is in vehicles. This number is converted from vehicles to feet based on assumed vehicle distance headway in the queue.

As a part of the IM program, road rangers provide free-of-charge services, such as tire changes, jumpstarts, and assisting in minor incidents to the motorists or highway patrol. The additional benefits provided by the service patrol are evaluated by estimating the number of activities by type performed by the service patrol and corresponding activity costs that are saved by road rangers. The default values for the rates of service patrol activities (per million vehicle miles traveled) and the activity costs are based on data for the FDOT District 4 service patrol program (*25*).

Dynamic Message Signs

The diversion rate due to disseminated traveler information should not be considered as constant and is independent of traffic and incident conditions. In reality, the number of DMS activations is a function of the number of incidents and lane blockage incidents, depending on the policy of DMS activation in the region. Below is a description of the methodology used in this study.

The default value for the number of DMS activations per year is assumed to equal the number of lane blockage incidents per year. This default value can be changed by the user. The rationale behind the use of this default value is that many agencies (including those in Florida) generally display incident information on the DMS only when there are lane blockage incidents.

This study assumes that the net diversion rate is a function of the potential saving in delay due to the incidents, as shown in Figure 1. If the calculated saving in delay is less than a certain minimum delay threshold (5 min or 20% is used as a default value in this study), then the diversion rate is assumed to be zero. This value is based on the findings reported by Mahmassani (*19*) and Peeta et al. (*21*) as discussed above. When the calculated saving in delay is high enough, the rate of diversion is assumed to be a maximum default value selected to be 40%, based on what is reported by Peeta et al. (*21*) and Khattak et al. (*22*). This maximum diversion rate can be changed by the user. Between the minimum delay and the maximum diversion rate thresholds, a linear relationship is assumed, as shown in Figure 1.

The total time savings resulting from DMS are calculated based on queuing equations (Equations 1 to 3 above) using reduced volumes due to diversions by travelers passing the DMS locations. Calculation of the benefits in this study accounts for the fact that the diverted vehicles will most likely experience additional travel time due to diverting to the alternative routes compared with the travel time on their original route without incident. This additional travel time was estimated and subtracted from the travel time benefits due to the reduction in queue length during incidents, which result from traffic diversion. To calculate this additional travel time, it was assumed that a certain proportion of the diverted travelers will divert to freeway segments and the remaining diverted travelers will divert to arterial streets. These proportions and the additional length of the alternative route will have to be estimated by the user of the methodology as they vary depending on local conditions. The average travel time on the alternative freeway and arterial routes is then calculated based on the average volume and average capacity of each of these two facility types, considering the additional traffic diverting to these routes due

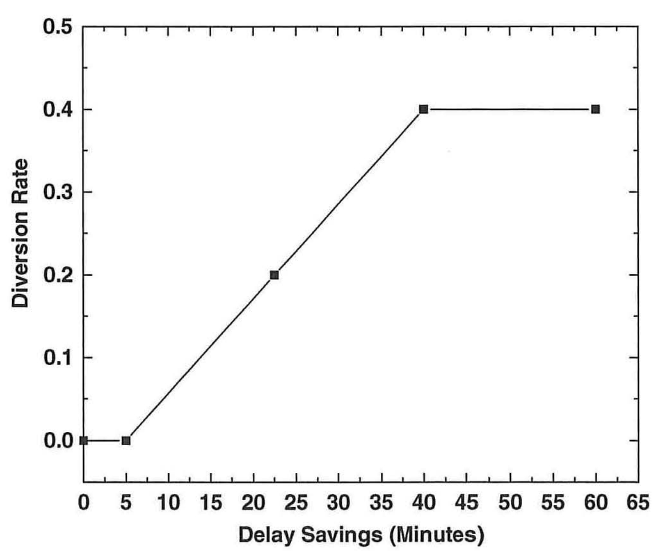


FIGURE 1 Relationship between diversion rate and expected delay savings.

to incidents. The additional time on the alternative route can then be calculated as the sum of the additional travel time on the freeway and arterial alternative routes.

Web- and Phone-Based ATIS

The main benefits of web-based and phone-based ATIS that can be converted to dollar values are assumed to result from changes in traveler trip decisions that are induced by the provided information. Instead of assuming fixed percentage reductions in time savings, the methodology proposed in this study explicitly calculates delay reductions based on the frequency of incident and duration. The delays without ATIS were calculated by using queuing theory, as explained in the IM and DMS sections above. The benefits of ATIS were then calculated by assuming that a certain percentage of traffic flow during incident conditions divert from the incident location and thus reducing the vehicle miles traveled at the incident location as follows:

$$q_{\text{ATIS}} = q_{\text{No\_ATIS}} \times (1 - \text{MP}) \times (1 - \text{CR}) \qquad (4)$$

where

q_{ATIS} and $q_{\text{No\_ATIS}}$ = volume of traffic joining the queue during incidents with and without ATIS, respectively;
MP = market penetration; and
CR = compliance rate.

MP can be assumed to vary between 1% and 10% depending on local data based on reviewing previous studies and data of 511 system calls in southeast Florida.

In addition, on the basis of studies reviewed above, the compliance rate was assumed to be a function of the estimated delay saving due to diversion, with 0% diversion for 0 min estimated saving in delay, 40% for 15-min saving in delay, 60% for 30-min saving in delay, and 100% for 45-min or more saving in delay. When using the developed tool and methodology, MP and CR may be updated to reflect local experiences with these systems and to account for new information that becomes available on the subject. Similar to the DMS, the fact that the diverted vehicles will most likely experience additional travel time on alternative routes is also taken into consideration.

Dollar Values of Benefits

The total annual benefits were converted into monetary savings by assigning corresponding dollar values to travel time, accidents, fuel costs, and emission costs of pollutants. Hadi et al. (*28*) examined the parameters that IDAS uses to convert various ITS impacts to dollar values and recommended changes to these parameters to reflect the values used in Florida. The suggested modified dollar values are used as default inputs in this study.

ITS Deployment Costs

An accurate estimation of the capital and recurrent (operation and management) costs is required for a reliable benefit–cost analysis of IM and ATIS. The annual costs of IM systems and DMS are derived based on the FDOT District 4 traffic management system costs. As shown in Table 1, the costs of IM consist of the cost for a traffic management center, closed-circuit television cameras, traffic detectors, fiber-optic communication lines, and road rangers.

An estimate of web- and phone-based ATIS costs should include the costs required for data gathering, data fusion and processing, telephonic dissemination platform, 511 code implementation, and a continuing marketing campaign (*13*). A number of reports were reviewed to determine the potential costs of these components (*29–34*). In this study, the average annual cost of ATIS is assumed to vary from $200,000 to $1,000,000, depending on the size and type of the system. This cost covers data fusion with traffic management center software and management of data from multiple sources: interactive voice recognition, website, telephone lines, and marketing. The 511 dialing cost is also added. It includes call routing, 511 translation, and toll-free translation. This cost is a function of the number of calls made to the system and their durations. When calculating this cost, the number of calls is estimated based on the total number of trips and market penetration.

TABLE 1 Estimate of Costs of IM and DMS

Component	Factors Affecting Costs	Capital Costs ($)	Operation and Maintenance Costs ($)
TMC	Size and type of TMC	3,000,000 (small area) 5,000,000 (medium area) 8,000,000 (large area)	500,000 (small area) 900,000 (medium area) 1,700,000 (large area)
CCTV camera assembly	Number of CCTV cameras	20,000	2,860
CCTV camera pole	Number of CCTV cameras	8,000	1,140
Traffic detector device	Number of detectors	7,000	255
Traffic detector pole	Number of detectors	4,000	145
Fiber-optic communication lines	Deployment length	30,000	600
Fiber-optic cable with junction boxes, splicing, termination, and conduit together with receivers and transceivers	Deployment length	100,000	2,000
Service patrol road ranger	Number of patrol vehicles, area of coverage, and vehicle shifts	0	43,250
DMS	Number of signs	105,000	10,000
DMS	Number of signs	105,000	10,000

NOTE: TMC = traffic management center; CCTV = closed-circuit television.

APPLICATIONS

The methodologies presented in the previous section were implemented in the FSUTMS, the traffic demand modeling environment in Florida, using the script language of the Cube software, which is the modeling engine of the FSUTMS. The FSUTMS model provides required inputs to the methodology, including traffic demands, link free-flow and loaded speeds, link capacities, and network geometry. This section presents the results from two case studies that were used to test the proposed methodology and to verify that the proposed evaluation methodology can be incorporated successfully in the travel demand forecasting environment.

The case studies are based on the network of Olympus, which is used in the FSUTMS Cube training course delivered by FDOT. The network includes 653 zones and 6,935 links. In Case Study 1, the IM system combined with eight DMS (four in each direction) are coded along a freeway segment. In Case Study 2, it is assumed that all the freeways and divided arterials are covered by web- and phone-based ATIS. The analysis is performed for the a.m. peak period, p.m. peak period, midday, and off-peak period and the resulting benefits are summed over these periods.

The annual benefit–cost analysis results for IM system and DMS are presented in Table 2. Although the results for each period are not shown because of space limitations, the benefits during the p.m. peak period (the most congested period) are much higher than in other periods, indicating that the IM is most effective under the congested condition. This finding can be explained by the fact that as the roadway becomes more congested—for example, during the p.m. peak period—more lane blockage incidents will occur, resulting in more diversions and thus greater reduction in the total incident response time and clearance time. The resulting benefit–cost ratio of 7.39 indicates that combined IM and DMS are cost-effective ITS alternatives for the Olympus network.

Tables 3 and 4 present the performance measures, annual benefits, and costs for web- and phone-based ATIS in Case Study 2. The major presented performance measures are the delay reduction and improvement in travel time reliability. Similar to IM, the time savings during the p.m. peak period is higher than those during other periods.

TABLE 3 Benefit–Cost Analysis Results for ATIS in Case Study 2

Year 2020 (MP = 7.5%)	Performance Measure	Dollar Value ($)
Time savings (vehicle hours)		
a.m. peak period	1,962.75	33,324.44
p.m. peak period	510,441.5	8,881,025.56
Midday	1,175.39	20,123.51
Off-peak period	121.12	2,090.68
Total	513,700.76	8,936,564.19
Total annual benefits	$8,936,564.19	
Total annual costs	$1,639,556.70	
Benefit–cost ratio	5.4	

NOTE: Annual costs = $300,000, dialing costs = 5 cents.

A further examination of the impacts of market penetration and annual costs is presented in Figure 2. As the number of ATIS users increases, the benefits of ATIS increase, but the rate of increase slows down for larger market penetrations because of the more congested traffic on the alternative routes. The cost of ATIS also increases with market penetration because of higher 511 dialing costs.

CONCLUSIONS

This study has presented an assessment of the methodologies used to assess the benefits and costs of IM and ATIS in existing sketch planning tools. On the basis of this assessment, improved methodologies were developed for this purpose. The improvements for IM evaluation compared with IM evaluation in existing tools include the incorporation of incident type, duration, and frequency in the delay calculations; consideration of secondary accident reduction; better estimation of fuel consumption and emissions based on queue length; and consideration of service patrol monetary benefits. The developed ATIS evaluation methodology can better estimate the number of DMS activations, ATIS market penetration, diversion rates based on incident and traffic conditions, and cost parameters of ATIS rather than based on fixed parameters as is done with other tools.

TABLE 2 Benefit–Cost Analysis Results for Scenario with Combined IM and DMS in Case Study 1

Year 2020	Performance Measure	Dollar Value ($)
Time savings (vehicle hours)	1,901,086.73	32,092,148.30
Safety (reduced total number of accidents)	23.76	4,109,926.06
Fuel savings (gallon)	2,176,714.26	6,236,279.70
CO emissions reduction (grams)	138.75	466,121.08
HC emissions reduction (grams)	15.78	26,141.16
NO_x emissions reduction (grams)	4.13	11,033.83
Road ranger (number of activities)	6,175	577,720.72
Total annual benefits	$43,519,370.85	
Capital costs	$1,924,267.65	
O&M costs	$3,966,326.50	
Total annual costs	$5,890,594.15	
Benefit–cost ratio	7.39	

NOTE: CO = carbon monoxide; HC = hydrocarbons; NO_x = oxides of nitrogen; O&M = operation and management.

TABLE 4 Benefit–Cost Analysis Results for ATIS in Case Study 2

Year 2020 (MP = 7.5%)	Performance Measure	Dollar Value ($)
Time savings (vehicle hours)		
a.m. peak period	1,962.75	33,324.44
p.m. peak period	510,441.5	8,881,025.56
Midday	1,175.39	20,123.51
Off-peak period	121.12	2,090.68
Total	513,700.76	8,936,564.19
Total annual benefits	$8,936,564.19	
Total annual costs	$2,979,113.40	
Benefit–cost ratio	3.00	

NOTE: Annual costs = $300,000, dialing costs = 10 cents.

The methodologies were implemented and integrated as part of the Florida demand modeling framework. The results from case studies that demonstrate application of the methodologies indicate that IM combined with DMS and phone-based and web-based ATIS deployments can be shown to be an effective solution from a benefit–cost point of view. The developed methodology and tool allow conducting various types of interesting and useful sensitivity analyses to justify the IM and ATIS deployment alternatives.

ACKNOWLEDGMENT

The work reported here was conducted as part of a project funded by the FDOT Research Center. The authors gratefully acknowledge the FDOT support of this effort.

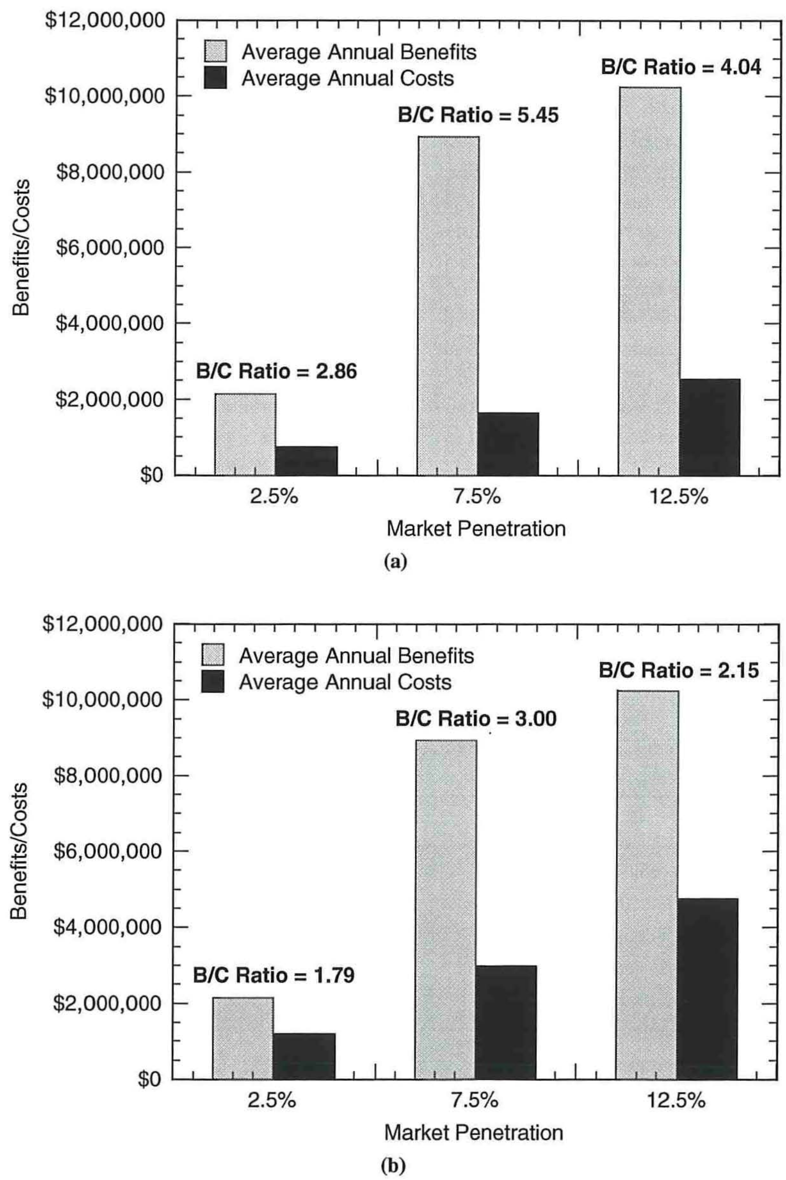


FIGURE 2 Sensitivity analysis results for ATIS market penetration in Case Study 2: (a) annual cost = $300,000, dialing cost = 5 cents, and (b) annual cost = $300,000, dialing cost = 10 cents; BC = benefit–cost.

REFERENCES

1. Science Applications International Corporation. *User's Manual for SCRITS, SCReening Analysis for ITS.* Prepared for Office of Traffic Management and ITS Applications, FHWA, 1999.
2. *IDAS-User's Manual.* Cambridge Systematics Inc., Cambridge, Mass., 2000.
3. Heither, C., and M. Thomas. *Testing of IDAS Capabilities Using Northeastern Illinois ITS Deployments.* Working paper 03-06. Chicago Area Transportation Study, Chicago, Ill., 2003.
4. May, A. D. *Traffic Flow Fundamentals.* Prentice Hall Inc., Englewood Cliffs, N.J., 1990.
5. Rakha, H. A., and W. Zhang. Consistency of Shock-Wave and Queuing Theory Procedures for Analysis of Roadway Bottlenecks. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C., 2005.
6. *Case Study 4: Mid-America Regional Council's (Kansas City) Enhanced Congestion Management System.* Cambridge Systematics Inc., Cambridge, Mass. http://idas.camsys.com/caseStudiesFrame.htm. Accessed June 4, 2009.
7. *Case Study 1: Ohio–Kentucky–Indiana Regional Council of Governments' Evaluation of ARTIMIS and ITS Program Plan.* Cambridge Systematics Inc., Cambridge, Mass. http://idas.camsys.com/caseStudies Frame.htm. Accessed June 4, 2009.
8. *Case Study 2: Michigan Department of Transportation Evaluation of the Temporary ITS for the Reconstruction of I-496 in Lansing, Michigan.* Cambridge Systematics Inc., Cambridge, Mass. http://idas.camsys.com/caseStudiesFrame.htm. Accessed June 4, 2009.
9. Kristof, T., M. Lowry, and G. S. Rutherford. *Assessing the Benefits of Traveler and Transportation Information Systems.* Project report. Washington State Department of Transportation, Olympia, 2005.
10. Yim, Y., and M. Miller. *Evaluation of TravInfo Field Operational Test.* California PATH Program, University of California, Berkeley, 2000.
11. Yim, Y., A. J. Khattak, and J. Raw. Traveler Response to New Dynamic Information Sources: Analyzing Corridor and Areawide Behavioral Surveys. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1803,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 66–75.
12. Peirce, S., and J. E. Lappin. Why Don't More People Use Advanced Traveler Information? Evidence from Seattle Area. Presented at 83rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2004.
13. *511 America's Traveler Information Number Deployment Assistance Report 1: Business Models and Cost Considerations.* 511 Deployment Coalition, FHWA, U.S. Department of Transportation, 2002.
14. Barfield, W., L. Conquest, J. Spyridakis, and M. Haselkorn. Information Requirements for Real-Time Motorist Information Systems. In *Proc. Vehicle Navigation and Information Systems Conference (VNIS),* New York, IEEE, 1989, pp. 101–112.
15. Madanat, S. M., C. Y. D. Yang, and Y.-M. Yen. Analysis of Stated Route Diversion Intentions under Advanced Traveler Information Systems Using Latent Variable Modeling. In *Transportation Research Record 1485,* TRB, National Research Council, Washington, D.C., 1995, pp. 10–17.
16. Benson, B. G. Motorist Attitudes about Content of Variable-Message Signs. In *Transportation Research Record 1550,* TRB, National Research Council, Washington, D.C., 1996, pp. 48–57.
17. Chatterjee, K., N. B. Hounsell, P. E. Firmin, and P. W. Bonsall. Driver Response to Variable Message Sign Information in London. *Transportation Research C,* Vol. 10, No. 2, 2002, pp. 149–169.
18. Tarry, S., and A. Graham. The Role of Evaluation in ATT Development. *Traffic Engineering and Control.* Vol. 36, No. 12, 1995, pp. 688–693.
19. Mahmassani, H. S. Dynamic Models of Commuter Behavior: Experimental Investigation and Application to the Analysis of Planned Traffic Disruptions. *Transportation Research A,* Vol. 24, No. 6, 1990, pp. 465–484.
20. Liu, Y. H., and H. S. Mahmassani. Dynamic Aspects of Commuter Decisions under Advanced Traveler Information Systems: Modeling Framework and Experimental Results. In *Transportation Research Record 1645,* TRB, National Research Council, Washington, D.C., 1998, pp. 111–119.
21. Peeta, S., J. L. Ramos, and R. Pasupathy. Content of Variable Message Signs and On-Line Driver Behavior. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1725,* TRB, National Research Council, Washington, D.C., 2000, pp. 102–108.
22. Khattak, A. J., F. C. Koppelman, and J. L. Schofer. Stated Preferences for Investigating Commuters' Diversion Propensity. *Transportation,* Vol. 20, 1993, pp. 107–127.
23. Huchingson, R. D., and C. L. Dudek. Delay, Time Saved, and Travel Time Information for Freeway Traffic Management. In *Transportation Research Record 722,* TRB, National Research Council, Washington, D.C., 1979, pp. 36–40.
24. Yang, S., F. Mei, and M. A. Hadi. Estimation of Incident Delays on Arterial Streets. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C., 2008.
25. Hadi, M., and C. Zhan. *Benefit–Cost Analysis of FDOT District 4 SMART SunGuide ITS Operations.* Final report. Prepared for Florida Department of Transportation, Tallahassee, 2006.
26. Zhu, X. Design Strategies for an Artificial Neural Network Based Algorithm for Automatic Incident Detection on Major Arterial Streets. Submitted to Institute of Transportation Engineers for the Daniel B. Fambro Student Paper Award Competition, 2008.
27. *Intelligent Transportation Systems 2005 Update Benefits, Costs, and Lessons Learned.* FHWA, U.S. Department of Transportation, 2005.
28. Hadi, M., P. Sinha, E. Quigley, and P. Hoke. *Intelligent Transportation Systems Deployment Analysis System Customization—Florida Specific Intelligent Transportation System Deployment Benefits.* Technical Memorandum 5. Prepared for Florida Department of Transportation, Tallahassee, 2004.
29. Sherer, E. *Preliminary Detail and Budget Estimate for ATIS and Telecommunications Components.* Prepared for Florida Department of Transportation, Tallahassee, 2003.
30. Raman, M., and C. Zimmerman. What Does It Cost to Deploy 511? A Case Study. Presented at 511 Deployment Coalition Meeting, Chicago, Ill., 2005.
31. Sanders, D., J. J. Sluss, J. P. Havlicek, and S. Radhakrishnan. *Deployment Plan—511 Traveler Information Services.* OTC Report 2005-1. Prepared for the Oklahoma Department of Transportation by the Oklahoma Transportation Center, Norman, 2005.
32. *511 Deployment Costs: A Case Study.* 511 Deployment Coalition, FHWA, 2006.
33. Berman, M. *Southeast Florida ATIS Management Plan.* Prepared for Florida Department of Transportation, Tallahassee, 2006.
34. Santana, R., et al. *The 511 South Florida Status Report.* Miami, 2007.

The Transportation Planning Applications Committee peer-reviewed this paper.

Accelerated Procedure of Multiclass Highway Traffic Assignment for Maryland Statewide Transportation Model

Xin Ye

The Maryland Statewide Transportation Model (MSTM) under development contains 20 user classes in the highway assignment procedure, including three types of autos: single-occupancy vehicles (SOV), high-occupancy vehicles (HOV) with two passengers, and HOV with three or more passengers classified by five income levels: commercial vehicle, median trucks, heavy trucks, regional autos, and regional trucks. A comprehensive disaggregation of user classes aids in capturing differences in how the transportation network is used. For example, SOV drivers are not allowed to use HOV lanes; trucks cannot be driven on truck-prohibited lanes; drivers at various income levels make different decisions about using toll roads because of their different values of time. However, a large number of user classes impose substantial computational burden on the procedure of multiclass highway assignment. If the conventional algorithm is used, the time consumption for this procedure is approximately proportional to the total number of user classes. In this study, a novel algorithm proposed by Robert Dial in 2006 is used to transfer min-path trees from one user class to another in the procedure of multiclass traffic assignment. This algorithm made it possible to speed up the highway assignment procedure of MSTM by a factor of 5.87.

Multiclass user-equilibrium (UE) assignment is the most prevalent approach to perform highway traffic assignment for statewide travel demand models in the United States (*1*). A typical statewide model has a large-scale network consisting of 100,000 to 250,000 links, thousands of zones, and a smaller number of user classes using different subsets of network links. For instance, vehicles may be classified as single-occupancy vehicles (SOV), high-occupancy vehicles (HOV), and trucks. SOV drivers are not allowed to use HOV lanes, which are exclusive for HOV, and trucks cannot be driven on truck-prohibited or HOV lanes. These vehicles may be further classified according to drivers' income level to capture their different tendency to use toll roads. For this purpose, the generalized cost for passing toll roads needs to vary across drivers at different income levels, and then the network cost is not entirely the same for all vehicles and drivers.

A comprehensive disaggregation of user classes aids in capturing vehicles' differences in using the transportation network but results in a dramatic increase in computational time for the procedure of network assignment. The conventional algorithm for UE assignment mainly consists of three steps: building up shortest paths, all-or-nothing (AON) assignment, and volume adjustment. Among these three steps, the most time-consuming part is usually the first step: building up shortest paths. Because of differences in link cost or restriction, shortest paths of different user classes cannot be entirely the same. If the conventional algorithm is used, shortest paths need to be individually generated for each specific user class. As a result, the time consumption for the procedure of multiclass UE assignment is approximately proportional to the number of user classes using different sets of shortest paths.

Dial (*2*) proposed a novel algorithm that allows modelers to transfer a set of shortest paths for one user class to that for another. The author tested this algorithm in a number of synthetic networks and found it could save a great amount of time in building up shortest paths for multiple user classes. As shortest path generation is the most time-consuming step in UE assignment, it should be promising to use this algorithm to accelerate the whole procedure of multiclass UE assignment for a statewide travel demand model. This study attempts to apply Dial's algorithm to a real statewide model: the Maryland Statewide Transportation Model (MSTM) to test whether and how many times this algorithm can speed up the procedure of UE assignment in a real multiclass transportation network.

REVIEW OF ALGORITHMS

This section reviews critical algorithms and briefly introduces how to apply them in this study. These algorithms include the conventional label-correcting algorithm for generating min-path trees, Dial's algorithm for transferring min-path trees, and the conventional Frank–Wolfe algorithm for UE assignment (*3*).

Basic Shortest Path Algorithm

The shortest path problem is famous in graph theory. Many algorithms have been invented to solve it by researchers from different fields such as mathematics, computer science, operations research, and transportation. In the transportation literature, Zhan and Noon (*4*) conducted a comprehensive evaluation of 15 algorithms in a variety of real transportation networks and recommended using the incremental graph algorithm implemented with two queues (*5*) to generate one-to-all shortest paths. By referring to that paper and the book by Ahuja et al. (*6*), the label-correcting algorithm implemented with a double-sided queue (*7*) was chosen as the basic algorithm for shortest path generation. The algorithm is coded as a function called the MinPathTree (see Appendix A).

National Center for Smart Growth Research and Education, 1112J, Preinkert Field House, University of Maryland, College Park, MD 20742. xinye@umd.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2176, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 84–91.
DOI: 10.3141/2176-10

Algorithm to Transfer Shortest Paths Between Two User Classes

Dial's algorithm (*2*) for a multiclass network is proposed based on the fact that travel cost difference among various user classes occurs on only a small number of links in a real network and most links have equal cost for all user classes. As a result, one user class's shortest paths should not differ substantially from those for another. One algorithm should help to save computational time if it allows one user class first to "inherit" shortest paths from another and then to generate new shortest paths by slightly modifying the old ones. This is the key point of Dial's algorithm, which allows for transferring a min-path tree from one user class X to another user class Y if all the link costs of user class Y are less than or equal to those of user class X. According to the book by Ahuja et al. (*6*), $\pi_i + c_{ij} \geq \pi_j$ for all link $ij \in A$ is a necessary and sufficient condition for a min-path tree, where π_i and π_j are potential costs to reach head and tail nodes of link ij, c_{ij} is link cost, and A is the link set. If a min-path tree has been generated for user class X, all these inequalities must be satisfied. If the cost on some links declines for user class Y, some relevant inequalities, $\pi_i + c_{ij} \geq \pi_j$, may not hold any more and their π_j (cost to reach node j) and pred_j (predecessor link of node j) values need to be updated. After π_j and pred_j are updated, relevant nodes j should be placed into the queue to restart the label-correcting procedure for satisfying all the inequalities.

The original paper mentions that "in the case where an arc (i.e., link) ij does not serve a class k, its cost c_{kij} is a constant M, where M is larger than the costliest path" (*2*, p. 852). In the coding procedure, it was found that using big M constant for restricted link cost considerably reduces the efficiency of the codes. For improving efficiency, a dummy "filter" array is defined to filter out links being restricted instead of using big M constant. Thus, there are two possible ways to reduce link cost in the procedure of min-path tree transfer: reducing cost or releasing restriction on some links. As per the suggestion from Dial (*2*), the identification (ID) of links was recorded with reduced cost or released restriction between two user classes into an array and those links with potential to change values in arrays "pred" and "pai" for predecessor link of each node and cost to reach each node were scanned. For better efficiency, the queue with nodes generated from the link scan is sorted according to the updated pai values in ascending order and then the sorted node ID is placed in the queue for the subsequent label correcting procedure. A function called MinPathTree_Transfer is coded to implement this algorithm (see details in Appendix B).

Algorithm for UE Assignment

The conventional Frank–Wolfe algorithm (*3*) is used to conduct multiclass UE assignment in this study. It is well known that the algorithm repeatedly executes three major steps: shortest path generation, AON assignment, and volume adjustment until the convergence criteria are satisfied. The novel algorithm to transfer min-path trees can save computational time only in the first step. However, as the first step consumes a large portion of time, a significant improvement can be expected in the speed of executing the entire algorithm.

INTRODUCTION TO MSTM AND ITS NETWORK

Overview of MSTM

MSTM is designed as a multilayer model working at both statewide and regional levels. The model contains 1,739 traffic analysis zones, including 1,607 state model zones (SMZs) and 132 regional model zones (RMZ). The 1,607 SMZs cover Maryland, Delaware, Washington, D.C., and parts of New Jersey, Pennsylvania, Virginia, and West Virginia. The 132 RMZs cover the United States, Canada, and Mexico. For SMZs at a statewide level, a traditional four-step travel demand model is developed to forecast passenger travel demand between origin–destination (O-D) pairs by various travel modes and time-of-day periods. The travel demand model contains

- Cross-classified models for production and attraction of 19 types of trips (seven trip purposes interactive with five travelers' income levels),
- Gravity models for distributing 19 types of trips into O-D trip matrices,
- Nested logit models for splitting O-D trip matrices into 11 travel modes (three auto modes—SOV, HOV with two passengers (HOV2), and HOV with three or more passengers (HOV3+)—and eight transit modes), and
- Time-of-day allocation model for splitting daily travel demand into demand over four daily time periods (a.m. peak, p.m. peak, midday, and night).

In addition to passengers' travel demand model, MSTM contains a freight demand model at the SMZ level:

- Cross-classified models for trip production and attraction of three types of vehicles (medium truck, heavy truck, and commercial vehicles) and
- Gravity models for distributing three types of truck trips into O-D trip matrices.

In MSTM, RMZs are designed for better measuring regional travel demand and its impact on the statewide modeling area. Similar to SMZs at a statewide level, MSTM has modules for forecasting regional passenger and freight demand but uses simpler methods:

- Regional O-D commodity flow data are initially obtained from a freight analysis framework (*9*) and then are converted to truck trips with conversion factors. For a finer spatial resolution, truck trips are disaggregated from trips between freight analysis framework regions to trips between counties as per employee distributions. Then, origins and destinations given by counties outside the SMZ area can be aggregated to RMZ, and those inside the SMZ area can be disaggregated to SMZ by employment.
- Regional passengers' travel demand is generated by Monte Carlo simulation based on survey data from the National Household Travel Survey 2001.

All modules mentioned above eventually generate 80 O-D trip matrices (1,739 × 1,739) for 20 user classes over four daily time periods. Twenty user classes include three types of autos (SOV, HOV2, HOV3+) by five income levels: commercial vehicle, median trucks, heavy trucks, regional autos, and regional trucks.

MSTM's Network for Highway Assignment

MSTM has both highway and transit networks but this paper focuses on the highway network. MSTM's highway network for the base year 2000 consists of 167,150 links relevant to 67,958 nodes. Within all the nodes, the initial 1,739 nodes represent zone centroids. As shown in Figure 1, the network covers the entire United States; it is denser in regions near the modeling area and sparser in those far

FIGURE 1 Overview of MSTM highway network.

away. Links far from the modeling area are composed of Interstates and highways. Highway assignment will be individually conducted by four time periods: a.m. peak (6:30 to 9:30 a.m.), midday (9:30 a.m. to 3:30 p.m.), p.m. peak (3:30 to 6:30 p.m.), and night (6:30 p.m. to 6:30 a.m.). Link cost, class, and restriction, which are critical for multiclass traffic assignment, are introduced in the following subsections.

Link Cost

In the procedure of highway assignment, the following function is formulated to calculate the generalized link cost for all user classes:

$$\text{cost} = t + (\text{toll}/\text{vot}) + 0.25 \times \text{distance} \tag{1}$$

where

t = link travel time (minutes), which is a function with respect to traffic volume assigned to the link;
toll = toll charge in cents, which differs in peak and off-peak periods;
vot = value of time (cents/min), converting toll charge in cents to time in minutes; and
distance = link distance in miles—the coefficient 0.25 converts distance in miles to time in minutes.

The values of time are given for auto drivers at five annual income levels ($0 to $20,000, $20,000 to $40,000, $40,000 to $60,000, 60,000 to $100,000, and $100,000+). vot takes values of 8.4, 25, 41.7, 50.4, and 106.4, denoted as vota, votb, votc, votd, and vote, respectively. vote, value of time for highest income level, is used for the five user classes: commercial vehicles, medium trucks, heavy trucks, regional trucks, and regional autos.

There are 64 toll links in the highway network, including a few ramps on US-267, a bridge on US-301, a few links representing the McHenry Tunnel and the Baltimore Harbor Tunnel, Baltimore Key Bridge on I-695, and some other links in Delaware. Toll charges range from $0.13 to $3.00.

Link Class

Links representing different types of transportation facilities should have different volume–delay relationships. Link classification helps modelers adopt different volume–delay functions for different kinds of links. In MSTM, the standard Bureau of Public Roads function (*10*) is formulated for the general volume–delay function:

$$t = t_0\left[1 + \alpha(v/c)^{\beta}\right] \tag{2}$$

where

α and β = coefficients that differ across link classes—two main link classes are defined for highway assignment, freeway and arterial, with α and β set up at 0.7 and 8 for freeway links and at 0.55 and 6 for arterial links, respectively;
t_0 = 60 × distance (mi)/free-flow speed (mph);
v = sum of total assigned traffic volumes of 20 user classes;
c = lane capacity (vehicles/h) × number of lanes/ConFac (conversion factor that adjusts road capacity to reflect better the traffic congestion in the traffic assignment);
ConFac = 0.39 for a.m. peak period (3 h), equivalent to expanding capacity by 2.56;
ConFac = 0.21 for midday peak period (6 h), equivalent to expanding capacity by 4.76;
ConFac = 0.34 for p.m. peak period (3 h), equivalent to expanding capacity by 2.94; and
ConFac = 0.22 for night period (12 h), equivalent to expanding capacity by 4.55.

As vehicles do not arrive in a uniform pattern during each time period, it is unreasonable to expand the capacity by the exact number of hours within that period. Instead, a certain level of deduction needs to be taken into account. For instance, in the a.m. peak period, the exact number of hours is 3 but the capacity is expanded by only 2.56 times to reflect traffic congestion during the a.m. peak period.

Link Restriction

Some highway links are restricted to some user classes. Sometimes links are restricted during the off-peak period and released in the peak period. Table 1 lists five link groups and the total number of links in each group by four time periods. Link Groups 1 and 2 represent nonhighway (e.g., rail-exclusive links) and vehicle-restricted highway links, which are always restricted to all highway user classes. Link Groups 3, 4, and 5 represent truck-prohibited, HOV2, and HOV3+

TABLE 1 Restricted Link Numbers by Time of Day

Link Group by Time Period	a.m.	p.m.	Midday	Night
Link Group 1 (nonhighway links)	1,082	1,082	1,082	1,082
Link Group 2 (vehicle-restricted links)	1,881	1,885	1,914	1,914
Link Group 3 (truck-prohibited links)	660	660	690	690
Link Group 4 (HOV2–only links)	156	154	0	0
Link Group 5 (HOV3+ only links)	73	71	0	0

links, respectively. Most truck-prohibited links are located in the region of Washington, D.C., including some lanes or roadway segments on MD-295, the George Washington Memorial Parkway, Arlington Boulevard in Northern Virginia near Washington, D.C., and Charles Street in Baltimore, Maryland. Most HOV2 and HOV3+ links represent I-95 HOV lanes, I-66 HOV lanes, and I-270 HOV lanes and some ramps on VA-267.

METHODOLOGY TO TRANSFER SHORTEST PATHS

Link Cost and Restriction by User Classes

Table 2 summarizes the definition of 20 user classes and their specific link restrictions and link costs. The initial 15 user classes are three types of auto users by five income levels. Because of their different values for toll charges, link costs for five income levels are differentiated as costa, costb, costc, costd, and coste, ranging from highest to the lowest. However, the cost difference occurs only on toll links, whereas non-toll links have the same cost for all user classes. Link costs for all the commercial vehicles, trucks, and regional autos are assumed to take the lowest values, coste, as it should be reasonable to assume that those users are the least sensitive to toll charges.

As for link restriction, medium trucks [User Class (UC) 17], heavy trucks (UC 18), and regional trucks (UC 19) are the most restricted user classes as they are not allowed to use either group of links, including nonhighway links [Link Group (LG) 1], vehicle-restricted links (LG 2), truck-prohibited links (LG 3), HOV2 links (LG 4), and HOV3+ links (LG 5). SOV users (UC 1 to 5) and commercial vehicle users (UC 16) are somewhat less restricted than truck users because they can use truck-prohibited links (LG 3). HOV2 users (UC 6 to 10) are more flexible than SOV users (UC 1 to 5) and commercial vehicle users (UC 16) because HOV2 lanes (LG 4) are available to them. HOV3+ users (UC 11 to 15) are the most flexible users because they can use all groups of links except nonhighway links (LG 1) and vehicle-restricted links (LG 2). Regional auto drivers (UC 20) are assumed to have the same level of flexibility as HOV3+ drivers (UC 11 to 15) have. Because auto occupancies are not differentiated for regional autos, all regional auto drivers are assumed able to use both kinds of HOV lanes (LG 4 and 5).

Alternative Ways to Transfer Shortest Paths

Because the algorithm can transfer one user class's shortest paths to those of another only when link cost is reduced or link restriction is released on some links, two alternative ways may be used to transfer min-path trees among 20 user classes, as shown in Figure 2. In the first way, priority may be given on cost reduction when transferring min-path trees. As shown in the upper part of Figure 2, an initial set of min-path tree, denoted as LG 1 to 5, costa, can be generated by calling the basic MinPathTree function with LG 1 to 5 restricted and costa used for link cost. This set of shortest paths is not used by any user classes but serves as a starting point for the whole procedure. After achieving the initial set, it is transferred to generate two new sets: LG 1, 2, 4, 5 costa and LG 1 to 5 coste by releasing LG 3 and reducing cost from costa to coste. The number in the arrow indicates that 660 links need to be scanned in the function MinPathTree_Transfer for generating the first new set and 64 links of toll roads need to be scanned for the second set. The first new set is used to conduct AON assignment for UC 1 and the second set is used for UC 17, 18, and 19, all of which share the same set of min-path tree.

Next, the set LG 1, 2, 4, 5 costa can be transferred to generate the set LG 1, 2, 5 costa by releasing LG 4. LG 4 consists of 156 links, all of

TABLE 2 Link Restriction and Cost by User Classes

User Class ID	User Classes	Link Restriction and Cost
1	SOV at Income Level 1	Excluding Link Group 1, 2, 4, 5; cost on toll links is the highest, denoted as "costa".
2	SOV at Income Level 2	Excluding Link Group 1, 2, 4, 5; cost on toll links is the 2nd highest, denoted as "costb".
3	SOV at Income Level 3	Excluding Link Group 1, 2, 4, 5; cost on toll links is the 3rd highest, denoted as "costc".
4	SOV at Income Level 4	Excluding Link Group 1, 2, 4, 5; cost on toll links is the 4th highest, denoted as "costd".
5	SOV at Income Level 5	Excluding Link Group 1, 2, 4, 5; cost on toll links is the lowest, denoted as "coste".
6	HOV-2 at Income Level 1	Excluding Link Group 1, 2, 5; cost on toll links is the highest ("costa").
7	HOV-2 at Income Level 2	Excluding Link Group 1, 2, 5; cost on toll links is the 2nd highest ("costb").
8	HOV-2 at Income Level 3	Excluding Link Group 1, 2, 5; cost on toll links is the 3rd highest ("costc").
9	HOV-2 at Income Level 4	Excluding Link Group 1, 2, 5; cost on toll links is the 4th highest ("costd").
10	HOV-2 at Income Level 5	Excluding Link Group 1, 2, 5; cost on toll links is the lowest ("coste").
11	HOV-3+ at Income Level 1	Excluding Link Group 1, 2; cost on toll links is the highest ("costa").
12	HOV-3+ at Income Level 2	Excluding Link Group 1, 2; cost on toll links is the 2nd highest ("costb").
13	HOV-3+ at Income Level 3	Excluding Link Group 1, 2; cost on toll links is the 3rd highest ("costc").
14	HOV-3+ at Income Level 4	Excluding Link Group 1, 2; cost on toll links is the 4th highest ("costd").
15	HOV-3+ at Income Level 5	Excluding Link Group 1, 2; cost on toll links is the lowest ("coste").
16	Commercial vehicles	Excluding Link Group 1, 2, 4, 5; cost on toll links is the lowest ("coste").
17	Medium trucks	Excluding Link Group 1–5; cost on toll links is the lowest ("coste").
18	Heavy trucks	Excluding Link Group 1–5; cost on toll links is the lowest ("coste").
19	Regional trucks	Excluding Link Group 1–5; cost on toll links is the lowest ("coste").
20	Regional autos	Excluding Link Group 1, 2; cost on toll links is the lowest ("coste").

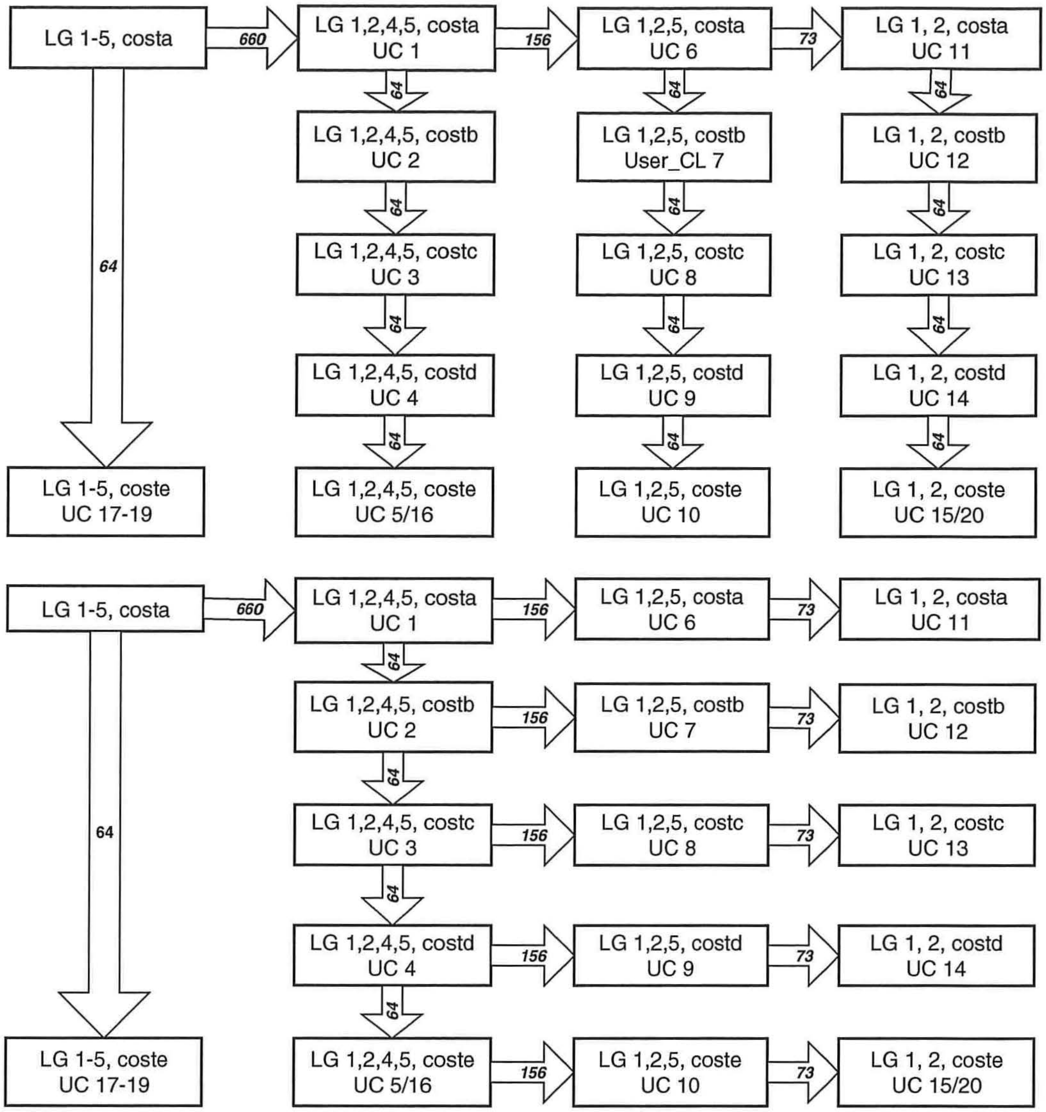


FIGURE 2 Two alternative ways to transfer shortest paths [LG: link group needs to be excluded in shortest path generation; UC: user class; the arrow represents the direction of shortest path transfer; the number in the arrow represents the number of links to be scanned in the algorithm (as in AM peak period)].

which need to be scanned in the function MinPathTree_Transfer. This new set of min-path tree will be used by UC 6. Then, this set can be transferred to generate set LG 1, 2 costa for UC 11 by releasing LG 5, which consists of 73 links. After achieving the sets LG 1, 2, 4, 5 costa; LG 1, 2, 5 costa; and LG 1, 2 costa they may be transferred to produce all other sets of min-path trees. During this procedure, link restrictions remain the same but link costs decline from costa → costb → costc → costd → coste in sequence. In each time of transfer, only 64 toll links need to be scanned to examine the potential cost reduction at their tail nodes. Through this procedure, 12 additional sets of min-path trees can be generated for all other user classes. UC 5 and 16 share the same set: LG 1, 2, 4, 5, coste and UC 15 and 20 share another same set: LG 1, 2, coste. It is a reasonable and efficient way to transfer and generate min-path trees for all user classes, but it is not the only way.

An alternative way is to transfer min-path trees with priority placed on restriction release. Its initial steps are the same as those in the first way but the differences start after the following four sets of min-path trees are achieved: LG 1, 2, 4, 5, costb; LG 1, 2, 4, 5, costc; LG 1, 2, 4, 5, costd; and LG 1, 2, 4, 5, coste. These four sets can be transferred to generate the remaining eight sets by sequentially releasing restrictions on LG 4 and 5, as illustrated in the lower part of Figure 2. The total number of links to be scanned in this way is a little greater than that in the first way. The second way may appear to be less efficient than the first. However, it is not the case in reality. Through comparison, it was found that the second way with priority on restriction release is about 15% faster than the first alternative in generating all sets of min-path trees. It indicates that experiments are required to test which transfer direction can create a more efficient way when alternative transfer directions are available.

Performance Evaluation

For performance evaluation, the new procedure is compared with the original one performed on Cube Voyager on a personal computer with an Intel Duo Core 2.1-GHz processor and 3.00 GB of random access memory. Cube Voyager was chosen as a benchmark because it is a mature commercial software for multiclass UE assignment. Before the formal test, the speed of Cube Voyager was compared with that of codes written in C programming language and they were found to perform almost equally fast in shortest path generation and AON assignment in the model. In the conventional procedure, 16 sets of min-path trees need to be generated for 20 user classes because some user classes share the same min-path tree. It was found that around 40 s is required to generate one set of shortest paths for all 1,739 zones, 1.3 s to conduct AON assignment for one trip table, and 15 s for volume adjustment. Thus, the total time consumption for each iteration is 681 s ($= 40 \times 16 + 1.3 \times 20 + 15$) in Cube Voyager. In the new procedure, as min-path trees can be transferred from one user class to another and need not be generated from the beginning, only 75 s is required to generate min-path trees for all user classes. These 75 s include the initial 40 s to generate the initial set of min-path tree not being used by any user classes and the remaining 35 s to generate the required 16 sets of min-path trees. On average, it takes only 2.2 s to generate each set of min-path tree, which is about 18 times faster than the conventional procedure. However, as the time required for AON assignment and volume adjustment in the new procedure does not differ from that in the conventional procedure, it still requires 116 s ($= 75 + 1.3 \times 20 + 15$) to complete one iteration in the new procedure. Finally, it turns out that MSTM's highway assignment procedure can be sped up by a factor of 5.87 (681/116) times with the new algorithm.

Figure 3 shows how highway assignment procedures converge over four time-of-day periods. Because the p.m. peak period is the most congested time period, it requires 54 iterations to reach relative gap at 0.01. The a.m. peak period is the second most congested period and 29 iterations are undertaken to reach that level of convergence. As midday and night periods are much less congested than peak periods, only eight and three iterations, respectively, are required to reach the same level of convergence. In total, 94 iterations are required to make four time periods reach acceptable convergence status. The total time assumption is around 3 h (116 s $\times$ 94/3,600 s). There are small variations in time consumption for different iterations but 116 s is the average value. If the conventional algorithm were performed in the same procedure, it would take almost 18 h.

CONCLUSIONS AND DISCUSSION

In this study, the novel algorithm proposed by Dial (*2*) was applied to MSTM's multiclass transportation network and the procedure of MSTM's highway assignment was accelerated by a factor of 5.87 compared with the conventional procedure in Cube Voyager. This application validates the conjecture in the original paper: "Test results using synthetic data suggest that its application to real networks should experience speedups of at least a factor of 2.0 and perhaps beyond 5.0." (*2*, p. 851). The effort undertaken in the study not only substantially improves the efficiency of the current MSTM but also enables modelers to pursue a more comprehensive and sophisticated modeling framework (e.g., adding interactive loops between land use and transportation models, traffic assignment on an hourly basis) that can be implemented within a reasonable time budget.

The limitation of this algorithm needs to be discussed. It can be used to transfer min-path trees efficiently only if cost reduction occurs on a small number of links. That is because a large number of links tend to result in too many nodes being placed in the queue. The QuickSort algorithm needs to be used to sort the queue according to nodes' potential values. It is well known that the QuickSort algorithm has average complexity as $O[N \times \log(N)]$ [i.e., the time consumption is proportional to $N \times \log(N)$, where N represents the number of elements in the array to be sorted], while the basic algorithm to generate a min-path tree is approximately an $O[N]$ algorithm in a real transportation network (i.e., the time consumption is proportional to N, where N represents the number of links of nodes in the network). (It may be debatable because the theoretical complexity of a label-correcting algorithm is nonlinear with respect to link number. However, as the real transportation network is rather sparse, the time consumption of the basic algorithm is almost proportional to the total number of links.) As a result, if the number of nodes in the queue goes beyond a certain threshold, the sorting procedure takes a longer time than generating the min-path tree from the beginning even if the time required for the subsequent label-correcting procedure is not included. In that case, it is not worth using this algorithm to transfer min-path trees because generating a new min-path tree from the beginning takes less

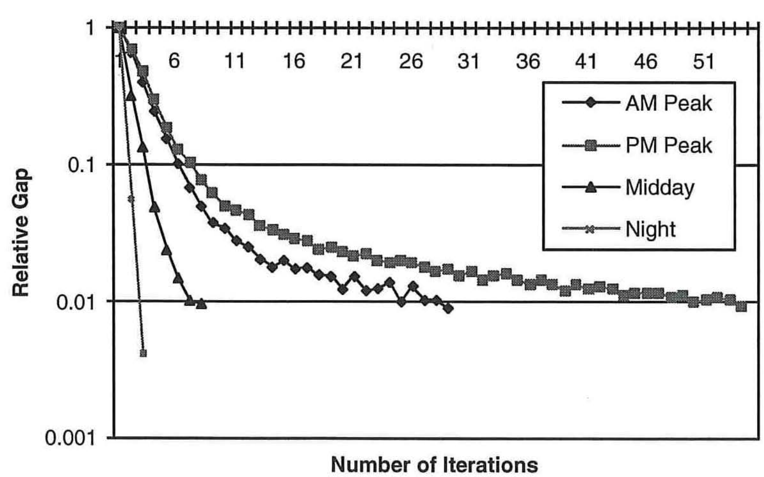


FIGURE 3 Convergence of highway assignments by time-of-day periods.

time. For the same reason, this algorithm may confront challenges in contemporary applications of traffic assignment such as turn penalties and traffic dynamics. With turn penalties, cars and trucks may receive different penalties at each node representing an intersection, which leads to much difference between cars' and trucks' link costs. Similarly, in dynamic traffic assignment, link travel time and intersection delay will be more accurately measured (e.g., through traffic simulation) for different user classes such as cars and trucks. In that case, as cost change frequently occurs on the network for different user classes, it will be difficult to apply Dial's algorithm to fasten the shortest path generation for different user classes. However, Dial's algorithm can substantially speed up a model with a large number of user classes whose cost differences occur on only a small portion of links, as evidenced by the MSTM. In the author's opinion, it is worthwhile to apply Dial's algorithm to save a great deal of computational time in similar modeling contexts.

ACKNOWLEDGMENTS

This research was funded by the Maryland Department of Transportation. The author is grateful to Gerrit Knaap of the National Center for Smart Growth Research and Education for his continued support. The author also thanks staff at Parsons Brinckerhoff, Inc. for their tremendous effort in network data preparation and model development.

APPENDIX A

BASIC MIN-PATH TREE ALGORITHM

MinPathTree(*A, c, o, N*_zones, *N*_nodes, filter):
Block links originating from all zones except those originating from the origin *o* (set up relevant filter values at 1);
Initialize five arrays, whose length is the total number of nodes:
pred[*i*] = −1; pai[*i*] = MAX_COST (e.g., 1.0e10); queue[*i*] = −1; in_queue[*i*] = 0; processed[*i*] = 0;
Define pointers "front" and "rear," initially pointing to the middle of array "queue";
Initialize pai[*o*] = 0.0, place the origin *o* into the front position of array queue and record status of node *o* as "in queue" but "not processed" (i.e., queue[front] = *o*, in_queue[*o*] = 1; processed[*o*] = 0);
Start iterations until the queue is empty (i.e., while (queue[front] ≠ −1)):
  Take the node *i* from the front position of the queue (i.e., *i* = queue[front]);
  Remove the node from the queue (i.e., queue[*i*] = −1);
  Record the node status as "not in queue" and "processed" (i.e., in_queue[*i*] = 0, processed[*i*] = 1);
  Move front pointer to the next position in the queue (i.e., front = front + 1);
  If front pointer goes beyond the bottom of the queue, move it to the top of the queue (i.e., if (front > *N*_node-1) front = front − *N*_node);
  Start iterations until all the links originating from node *i* have been scanned:
    If the link is blocked (i.e., the corresponding filter value is 1), scan the next link;
    Let *j* take the tail node of the current link;
    Let *t* take the travel cost on the current link (i.e., *t* = *c*[the current link ID]);
    If (pai[*i*] + *t* < pai[*j*]), then pai[*j*] = pai[*i*] + *t* and pred[*j*] = the current link ID;
    If the node *j* is not in the queue (i.e., in_queue[*j*] = 0),
      If the node *j* has not been processed before (i.e., processed[*j*] = 0),
      move "rear" pointer to the next position of the queue (i.e., rear = rear + 1).
      If the rear pointer goes beyond the bottom of the queue array, move it to the top of the queue (i.e., if (rear > *N*_node-1) rear = rear − *N*_nodes).
      Place node *j* in the rear position of the queue (queue[rear] = *j*).
      Record the status of node *j* as in queue (in_queue[*j*] = 1).
    Else,
      move front pointer to the previous position of the queue (i.e., front = front − 1).
      If front pointer goes beyond the top of the queue array, move it to the bottom of the queue (i.e., if (front < 0) front = front + *N*_nodes).
      Place node *j* in the front position of the queue (queue[front] = *j*).
      Record the status of node *j* as in queue (in_queue[*j*] = 1).
Return (pred, pai)
END

Note: There are six sets of arguments in this function: *A* represents network topology; *c* represents cost for each link in the network; *o* is the ID of the original zone; *N*_zones and *N*_nodes represent the total number of zones and nodes in the network, filter represents arrays indicating whether the link is blocked. Finally, the function returns two arrays: pred and pai. pred array records the predecessor link ID of each node on shortest paths. A shortest path from origin to destination can be retrieved by tracing back through predecessor links on the min-path tree while AON assignment can be performed. pai array records accumulated travel cost from the origin to each node (also called potential). If the pai value remains at the initial value of MAX_COST after the function is implemented, the corresponding node is not accessible from the origin.

APPENDIX B

ALGORITHM TO TRANSFER MIN-PATH TREE

MinPathTree_Transfer(old_pai, old_pred, link_ID_set, *A, c, o, N*_zones, *N*_nodes, filter):
Initialize two arrays "pai" and "pred" and copy "old_pai" and "old_pred" arrays to them.
Scan all the links in link_ID_set:
  Let *k*1 and *k*2 take head node and tail node of the current link;
  Let *t* take the travel cost on the current link;
  If (pai[*k*1] + *t* < pai[*k*2]),
    pai[*k*2] = pai[*k*1] + *t*;
    pred[*k*2] = the current link ID;
    Record node *k*2 and pai[*k*2] into a table called "node_pai."
Use QuickSort algorithm (*8*) to sort the table node_pai according to pai values in ascending order;
Place the sorted node IDs into a queue and record the status of those nodes as "in queue" but "not processed";
Define pointers "front" and "rear," respectively, pointing to the front and the rear positions of the queue with node IDs;

(The rest of them are the same as those in function MinPathTree( ).)
Start iterations until the queue is empty:
  Take the node i from the front position of the queue;
  Remove the node from the queue;
  Record the node status as "not in queue" and "processed";
  Move front pointer to the next position in the queue;
  If front pointer goes beyond the bottom of the queue, move it to the top of the queue;
  Start iterations until all the links originating from node i have been scanned:
    If the link is blocked, scan the next link;
    Let j take the tail node of the current link;
    Let t take the travel cost on the current link;
    If (pai[i] + t < pai[j]), then pai[j] = pai[i] + t and pred[j] = the current link ID;
      If the node j is not in the queue,
        If the node j has not been processed before,
          move rear pointer to the next position of the queue.
          If the rear pointer goes beyond the bottom of the queue array, move it to the top of the queue.
          Place node j in the rear position of the queue.
          Record the status of node j as in queue.
        Else,
          move front pointer to the previous position of the queue
          If front pointer goes beyond the top of the queue array, move it to the bottom of the queue.
          Place node j in the front position of the queue.
          Record the status of node j as in queue.
Return (pred, pai)
END

Note: There are nine sets of arguments in this function, most of which are the same as those in the function MinPathTree except the first three arrays: old_pai, old_pred, and link_ID_set. The first two arrays pass the information of the previous user class min-path tree to the function and the third array passes IDs of links to be scanned. Arrays c and filter carry link cost and restriction information for the current user class, rather than for the previous one. The function returns two arrays: pred and pai for the current user class.

REFERENCES

1. Horowitz, A. *NCHRP Synthesis of Highway Practice 358: Statewide Travel Forecasting Models.* Transportation Research Board of the National Academies, Washington, D.C., 2006.
2. Dial, R. B. An Efficient Algorithm for Building Min-path Trees for All Origins in a Multi-class Network. *Transportation Research B,* Vol. 40, No. 10, 2006, pp. 851–856.
3. Frank, M., and P. Wolfe. An Algorithm for Quadratic Programming. *Naval Research Logistics Quarterly,* Vol. 3, 1956, pp. 95–110.
4. Zhan, F. B., and C. E. Noon. Shortest Path Algorithms: An Evaluation Using Real Road Networks. *Transportation Science,* Vol. 32, No. 1, 1998, pp. 65–73.
5. Pallottino, S. Shortest-Path Methods: Complexity, Interrelations, and New Propositions, *Networks,* Vol. 14, 1984, pp. 257–267.
6. Ahuja, R. K., T. L. Magnanti and J. B. Orlin. *Network Flows: Theory, Algorithms and Applications,* Prentice Hall, Englewood Cliffs, N.J., 1993.
7. Pape, U. Implementation and Efficiency of Moore Algorithms for the Shortest Root Problem. *Mathematical Program,* Vol. 7, 1974, pp. 212–222.
8. Hoare, C. A. R. Algorithm 61 (Partition) and Algorithm 65 (Find). *Communications of the ACM,* Vol. 4, No. 7, 1961, pp. 321–322.
9. *Freight Analysis Framework (FAF).* FHWA. http://ops.fhwa.dot.gov/freight/freight_analysis/faf/index.htm. Accessed June 30, 2008.
10. *Traffic Assignment Manual.* Bureau of Public Roads, U.S. Department of Commerce, Urban Planning Division, Washington D.C., 1964.

The Statewide Multimodal Transportation Planning Committee peer-reviewed this paper.

TRANSPORTATION RESEARCH RECORD: JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Peer Review Process

The *Transportation Research Record: Journal of the Transportation Research Board* publishes approximately 25% of the more than 3,000 papers that are peer reviewed each year. The mission of the Transportation Research Board (TRB) is to disseminate research results to the transportation community. The Record series contains applied and theoretical research results as well as papers on research implementation.

The TRB peer review process for the publication of papers allows a minimum of 30 days for initial review and 60 days for rereview, to ensure that only the highest-quality papers are published. At least three reviews must support a committee's recommendation for publication. The process also allows for scholarly discussion of any paper scheduled for publication, along with an author-prepared closure.

The basic elements of the rigorous peer review of papers submitted to TRB for publication are described below.

Paper Submittal: June 1–August 1

Papers may be submitted to TRB at any time. However, most authors use the TRB web-based electronic submission process available between June 1 and August 1, for publication in the following year's Record series.

Initial Review: August 15–November 15

TRB staff assigns each paper by technical content to a committee that administers the peer review. The committee chair assigns at least three knowledgeable reviewers to each paper. The initial review is completed by mid-September.

By October 1, committee chairs make a preliminary recommendation, placing each paper in one of the following categories:

1. Publish as submitted or with minor revisions,
2. Publish pending author changes and rereview, or
3. Reject for publication.

By late October, TRB communicates the results of the initial review to the corresponding author indicated on the paper submission form. Corresponding authors communicate the information to coauthors. Authors of papers in Category 2 (above) must submit a revised version addressing all reviewer comments and must include a cover letter explaining how the concerns have been addressed.

Rereview: November 20–January 25

The committee chair reviews revised papers in Category 1 (above) to ensure that the changes are made and sends the Category 2 revised papers to the initial reviewers for rereview. After rereview, the chairs make the final recommendation on papers in Categories 1 and 2. If the paper has been revised to the committee's satisfaction, the chair will recommend publication. The chair communicates the results of the rereview to the authors.

Discussions and Closures: February 1–May 15

Discussions may be submitted for papers that will be published. TRB policy is to publish the paper, the discussion, and the author's closure in the same Record.

Many papers considered for publication in the *Transportation Research Record* are also considered for presentation at TRB meetings. Individuals interested in submitting a discussion of any paper presented at a TRB meeting must notify TRB no later than February 1. If the paper has been recommended for publication in the *Transportation Research Record*, the discussion must be submitted to TRB no later than April 15. A copy of this communication is sent to the author and the committee chair.

The committee chair reviews the discussion for appropriateness and asks the author to prepare a closure to be submitted to TRB by May 15. The committee chair reviews the closure for appropriateness. After the committee chair approves both discussion and closure, the paper, the discussion, and the closure are included for publication together in the same Record.

Final Manuscript Submittal: March 15

In early February, TRB requests a final manuscript for publication—to be submitted by March 15—or informs the author that the paper has not been accepted for publication. All accepted papers are published by December 31.

Paper Awards: April to January

The TRB Executive Committee has authorized annual awards sponsored by Groups in the Technical Activities Division for outstanding published papers:

- Charley V. Wootan Award (Policy and Organization Group);
- Pyke Johnson Award (Planning and Environment Group);
- K. B. Woods Award (Design and Construction Group);
- Patricia F. Waller Award (Safety and System Users Group);
- D. Grant Mickle Award (Operations and Preservation Group); and
- John C. Vance Award (Legal Resources Group).

Other Groups also may nominate published papers for any of the awards above. In addition, each Group may present a Fred Burggraf Award to authors 35 years of age or younger.

Peer reviewers are asked to identify papers worthy of award consideration. Each Group reviews all papers nominated for awards and makes a recommendation to TRB by September 1. TRB notifies winners of the awards, which are presented at the following TRB Annual Meeting.

Transportation Research Board
www.TRB.org

ISBN 978-0-309-16045-2
90000
9 780309 160452